DOCUMENT RESUME

ED 249 215 SP 025 343

AUTHOR Wood, Frank H.
TITLE Formal Observation of Students' Social Behavior.
INSTITUTION Minnesota Univ., Minneapolis. National Support Systems Project.
SPONS AGENCY Office of Special Education and Rehabilitative Services (ED), Washington, DC. Div. of Personnel Preparation.
PUB DATE Dec 80
GRANT OEG007902045
NOTE 96p.; For other modules in this series, see ED 238 844 and SP 025 332-354. For the genesis of these modules, see ED 186 399. Reviewed by K. Charlie Lakin.
PUB TYPE Guides - Classroom Use - Guides (For Teachers) (052)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Classroom Observation Techniques; Higher Education; Interpersonal Competence; Learning Modules; *Mainstreaming; Preservice Teacher Education; *Social Behavior; Student Behavior; *Teacher Education Curriculum; *Teacher Educators
IDENTIFIERS Education for All Handicapped Children Act

ABSTRACT

This module (part of a series of 24 modules) is on teachers' use of systematic observation records of social behavior to aid in assessing students' special needs and in evaluating the effects of specific programs. The genesis of these materials is in the 10 "clusters of capabilities," outlined in the paper, "A Common Body of Practice for Teachers: The Challenge of Public Law 94-142 to Teacher Education." These clusters form the proposed core of professional knowledge needed by teachers in the future. The module is to be used by teacher educators to reexamine and enhance their current practice in preparing classroom teachers to work competently and comfortably with children who have a wide range of individual needs. The module includes objectives, scales for assessing the degree to which the identified knowledge and practices are prevalent in an existing teacher education program, and self-assessment test items. Journal articles are appended expanding on the topic of formal observation of students' social behavior. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED249215

FORMAL OBSERVATION OF STUDENTS' SOCIAL BEHAVIOR

PREPARED BY
FRANK H. WOOD

REVIEWED BY
K. CHARLIE LAKIN

DECEMBER 1980

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

SP 025 343
SP025 343

This paper presents one module in a series of resource materials which are designed for use by teacher educators. The genesis for these materials is in the ten "clusters of capabilities," outlined in the paper, "A Common Body of Practice for Teachers: The Challenge of Public Law 94-142 to Teacher Education", which form the proposed core of professional knowledge needed by professional teachers who will practice in the world of tomorrow. The resource materials are to be used by teacher educators to reexamine and enhance their current practice in preparing classroom teachers to work competently and comfortably with children who have a wide range of individual needs. Each module provides further elaboration of a specified "cluster of capabilities" - in this case, Referral: Formal observation of students' social behavior.

Extending the Challenge:

Working Toward a Common Body of Practice for Teachers

Concerned educators have always wrestled with issues of excellence and professional development. It is argued, in the paper "A Common Body of Practice for Teachers: The Challenge of Public Law 94-142 to Teacher Education," that the Education for All Handicapped Children Act of 1975 provides the necessary impetus for a concerted reexamination of teacher education. Further, it is argued that this reexamination should enhance the process of establishing a body of knowledge common to the members of the teaching profession. The paper continues, then, by outlining clusters of capabilities that may be included in the common body of knowledge. These clusters of capabilities provide the basis for the following materials.

The materials are oriented toward assessment and development. First, the various components, rating scales, self-assessments, sets of objectives, and respective rationale and knowledge bases are designed to enable teacher educators to assess current practice relative to the knowledge, skills, and commitments outlined in the aforementioned paper. The assessment is conducted not necessarily to determine the worthiness of a program or practice, but rather to reexamine current practice in order to articulate essential common elements of teacher education. In effect then, the "challenge" paper and the ensuing materials incite further discussion regarding a common body of practice for teachers.

Second and closely aligned to assessment is the developmental perspective offered by these materials. The assessment process allows the user to view current practice on a developmental continuum. Therefore, desired or more appropriate practice is readily identifiable. On another,

perhaps more important dimension, the "challenge" paper and these materials focus discussion on preservice teacher education. In making decisions regarding a common body of practice it is essential that specific knowledge, skill and commitment be acquired at the preservice level. It is also essential that other additional specific knowledge, skill, and commitment be acquired as a teacher is inducted into the profession and matures with years of experience. Differentiating among these levels of professional development is paramount. These materials can be used in forums in which focused discussion will explicate better the necessary elements of preservice teacher education. This explication will then allow more productive discourse on the necessary capabilities of beginning teachers and the necessary capabilities of experienced teachers.

In brief, this work is an effort to capitalize on the creative ferment of the teaching profession in striving toward excellence and professional development. The work is to be viewed as evolutionary and formative. Contributions from our colleagues are heartily welcomed.

FORMAL OBSERVATION OF STUDENTS' SOCIAL BEHAVIOR

This module reflects the author's experience that, while many classroom teachers respond enthusiastically to new, and often helpful, educational perspectives and procedures, systematic observation records of social behavior add both to their ability to assess students' special needs and to evaluate the effects of specific programs. Often, however, teachers are prevented by time and setting constraints from using on a regular basis the observation procedures they are taught in some training programs. In many cases, instead of learning principles for the flexible application of observation procedures, they have been given practice with specific observation procedures that are too complicated and time consuming for regular classroom use. This module is intended to guide teacher trainers in exposing teachers to a sampling of simple observation and recording procedures so that they may learn to view systematic observation as a tool to be used when needed rather than a burden to be avoided unless required as a course assignment. The procedures suggested in this module can be remembered and used to good effect in the classroom with little preparation or practice. The discussion focuses on issues in the choice of the right procedure to meet a specific classroom need. Although some of the procedures described could not be recommended for use in research, there will be substantial gains in professional practice if teachers learn to apply them. To this end, content has been arranged to provide the teacher educator with materials for introducing this important topic to pre- and in-service teachers through a lecture/discussion approach.

The resulting treatment of observation procedures will not satisfy the needs of the researcher who wishes to use observation for a scientific study of human behavior. Nor will it meet the needs of the psychologist or teacher educator who wishes to develop a solid understanding of the principles underlying the application of observation methodology to education. For these

users, additional readings are suggested, and two representative papers have been included as an appendix.

Other modules in the total set which are related to this topic include:

Class Management	Module 3A
Developing Goals and Objectives for IEPs	Module 9A

Contents

Within this module are the following components:

Set of Objectives - The objectives focus on the teacher educator rather than as a student (preservice teacher). They identify what can be expected as a result of working through the materials. The objectives which apply to teachers are also identified. They are statements about skills, knowledge, and attitudes which should be part of the "common body of practice" of all teachers. Page 1

Rating Scales - Scales are included by which a teacher educator could, in a cursory way, assess the degree to which the knowledge and practices identified in this module are prevalent in the existing teacher-training program. The rating scales also provide a catalyst for further thinking in each area. Page 2

Self-Assessment - Specific test items were developed to determine a user's working knowledge of the major concepts and principles in each subtopic. The self-assessment may be used as a pre-assessment to determine whether one would find it worthwhile to go through the module or as a self check, after the materials have been worked through. The self-assessment items also can serve as examples of mastery test questions for students. Page 3

Rationale and Knowledge Base - This section summarizes a number of methods for systematically observing and recording the social behavior of students on class management. The more salient concepts and strategies are reviewed. A few brief observation activities for students are outlined and attached at the end of this section. Page 8

Bibliography - A partial bibliography of useful books and materials is included after the list of references. Page 55

Articles - Two articles (reproduced with author's permission) accompany the aforementioned components. The articles support and expand on the knowledge base. Page 56

Objectives for Teacher Educators and for Incorporation into Teacher Education Curriculum

1. To distinguish between student behavior that can be observed directly and student problems which can only be inferred from directly observed behavior.

2. To understand the role played by observations of behavior and inferences based on them in making decisions about students' special needs.

3. To understand the practical implications of the fact that no behavior record is a complete record of all that has occurred in a given setting.

4. To know the differences between the following observation procedures and their appropriate uses to record the behavior of individual students or groups:

    a. anecdotal records of critical incidents.
    b. narrative records.
    c. behavior logs.
    d. behavior rating scales or checklists.
    e. behavior objectives sequences.
    f. observation schedules.
    g. event frequency counts.
    h. summary charts and graphs.

5. To understand how to keep simple records of interactions among individuals in a group.

6. To understand how to record the collective behavior of an entire group of individuals.

7. To understand the practical implication of the concepts of reliability and validity as applied to observation records.

8. To apply several observation recording procedures in actual or videotaped situations.

Rating Scale for Teacher Preparation Programs

____1. Students being prepared for teaching do not use observation as a source of information about pupils' behavior, placing major reliance on standardized testing procedures. Observation recording is sporadic and of the anecdotal or critical incidence variety. Stress is on observation of students who attract attention by their "acting-out" behavior.

____2. Students practice writing narrative records or summary reports as part of the assessment and evaluation of pupil behavior. Observation limited to unsystematic approaches. Stress is on observation of students who attract attention by their "acting out" behavior.

____3. Students use behavior checklists in addition to narrative records as a means of recording observations of student behavior. Stress still on observation of students who attract attention by the "acting-out" behavior.

____4. Students know how to do event frequency recording as a means of measuring changes over time and use this in addition to narrative records and checklists. The focus still tends to be on the observation of students who attract attention by "acting-out" behavior.

____5. Students learn a variety of observation/recording procedures together with their uses. Procedures are used routinely for assessment, monitoring and evaluation of pupil behavior and changes in behavior over time. Observations are made of all students and individual records are interpreted in the context of the behavior of others in the group, including that of the teacher.

Self-Assessment

1. People act in all of the ways listed below. Put an "O" in the space in front of those that can be observed directly and a "I" in front of those that must be inferred from other actions that can be directly observed:

____Student hits another student
____Student sits down in a chair.
____Student feels angry.
____Student expresses affection for her teacher.
____Student writes on a piece of paper.
____Student throws a paper airplane.
____Student puts his hand on the teacher's shoulder.
____Student hates the principal.

2. Inferences about the feelings that accompany behavior should be recorded:

____a. never.
____b. only when they are clearly expressed.
____c. only when the behavior on which they are based is also recorded.
____d. both b. and c.

3. List at least three ways to make observation recording easier for the observer.

a.

b.

c.

4. Incomplete observation records are necessarily less_____ than more complete observation records.

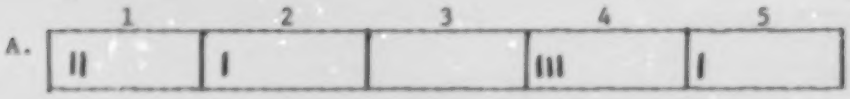
____a. reliable.
____b. valid.
____c. both a. and c.

5. How does an "anecdotal record" of a student's behavior during one 15 minute session differ from a "narrative record" of that same behavior.

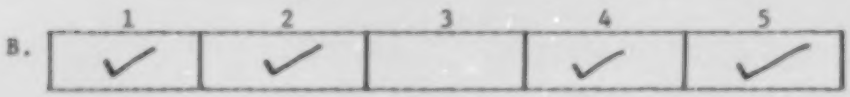
____a. it is less objective.
____b. it is less complete.
____c. it is written down after the behavior occurs.
____d. it is more easily recorded by an observer.

6. You wish to obtain a parent's view of his son's behavior during the coming week. What method of observation recording would you use? Why?

7. Below are examples of interval and frequency recording of a student's talking out without permission during a five minute observation period. Label each.



_____



_____

Which is the more complete record of talking out without permission?
Which would be easier for a teacher to keep while he/she is teaching?

8. The following are items chosen from a problem behavior checklist. Mark the space in front of those which require the rater to make inferences about a student's behavior in responding.

____a. doesn't speak

____b. speech unintelligible

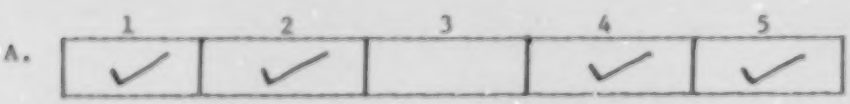
____c. destructive of property

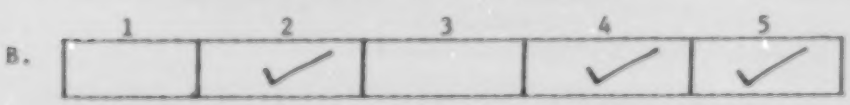
____d. takes things belonging to others without asking them

____e. irresponsible

____f. plays with younger children

9. Two observers observed the same students at work for five minutes. At the end of each minute, they mark on their record if the student has been attending to his/her work. Here are their records:





What can you say about the reliability of their observation records?

9. (continued)

What can you say about the validity of their observation records?

10. Here are two descriptive statements about a student.

a. Student hits his classmates frequently.

b. Student expresses deep-seated feelings of anger frequently.

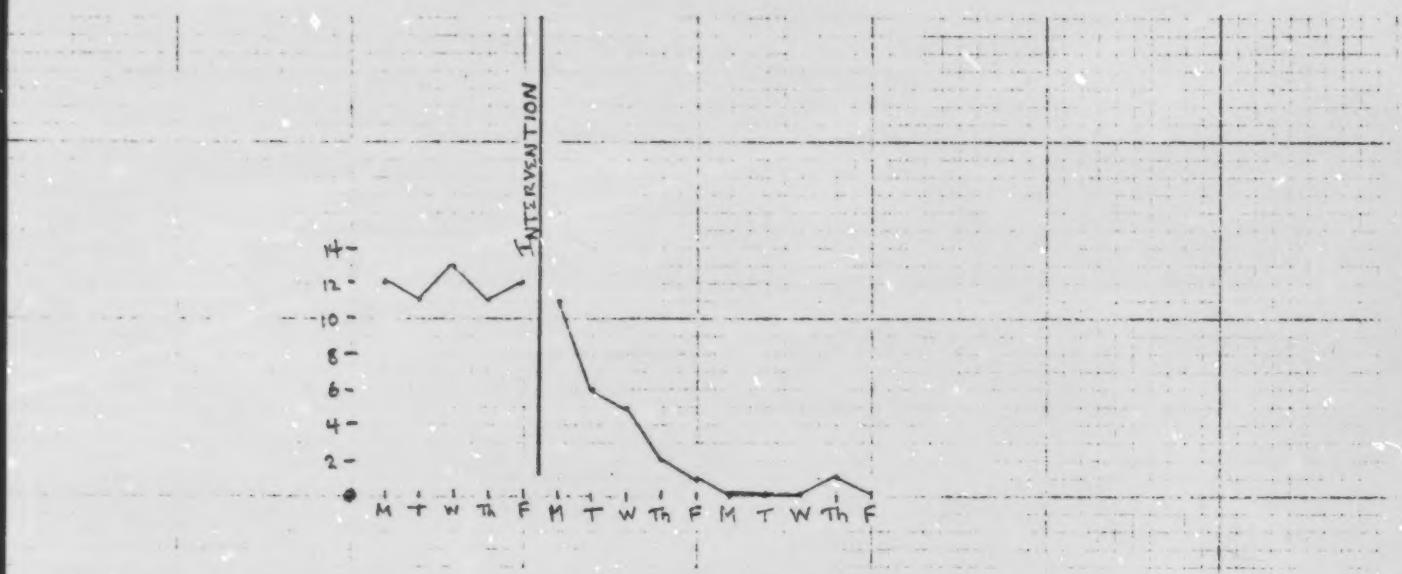
For which of these two statements will it be easier to obtain a reliable observation record?

_____

11. How are a behavior rating scale and a behavioral objectives sequence alike? How are they different?

12. Behavior ratings are often compared to average ratings furnished by the developer of the rating scale. What other comparisons would it be important to make in evaluating the "normality" of observed behavior? How would you obtain the needed information?

13. A teacher has kept data on the total number of questions asked by her students after she introduces the daily math lessons. Because her original count showed too many questions were being asked, she introduced an intervention to bring the number of questions down. Feeling that many students were simply finding it easier to ask questions than read the directions, she informed the class that all questions were to be written on a piece of paper so as not to disturb those who were already beginning to work. Here is her observation record for a three week period. Was this a good choice of the behavior to observe? To seek to decrease? What does the record show? What questions would you have asked the teacher? What other information might you want to collect?



14. Here is a sample narrative record. Encircle the parts of the record that required inference by the observer.

Chuck is seated at his desk, smiling. The teacher, in starting a discussion about the character of Hester Prynne in Hawthorne's "Scarlet Letter," asks members of the class to volunteer their ideas. Chuck responds with a loud laugh and the remark, "She's a whore." This draws laughter from some less mature members of the class. As usual, he is seeking ways to disrupt the discussion. The teacher continues with the question, "Can we analyze her motivation?" Jeff offers to answer. The teacher says, "Yes, Jeff." At this point, most students are looking at Jeff as he begins his answer, "I read it several times, and ...." Chuck tries again to attract attention by blurting out, "I only read it once." Most of the class would definitely like to hear what Jeff has to say without having these interruptions, but there is some laughter again. Jeff continues, "I think she is basically introverted." Chuck says loudly, "Introverted, she's not introverted. She's perverted." Many students laugh in embarrassment at his crude remark. The teacher tries to make Chuck be serious by saying to him, "Let's go beyond that...." Chuck shakes his head, pretending to be embarrassed. "You said she was perverted," says the teacher, "Now, back it up from the book." Chuck replies, "you know what they were doing out there in the woods, her and her imp kid. I could just read between the lines. My mother wouldn't want me to know about those things." Of course, this really just illustrates what a dirty mind Chuck has.

Note: Adapted from material from the CONSERT Project, by P. Pattavina and E. A. Gotts, School of Human Development, University of Texas at Dallas. 1979.

15. "Most teachers who observe and record systematically accept the behaviorist view of instruction and learning. Records of that kind are not very important to those of us who place importance on the development of good feelings about self and others in our classrooms." Comment on this statement.

Formal Observation of Students' Social Behavior

Special educators who are responsible for direct service to students with educational disabilities have found systematic methods of observing and recording relevant student behavior to be a valuable tool for aiding their students' learning. Application of these methods requires additional teacher time, which suggests that those teachers who use them must be repaid by an increased sense of professional competence and personal satisfaction. The teacher willing to commit this additional time, which is greater during the initial learning period, will find that observation methods are easily mastered and the rewards are real. Observation data alone are not an adequate basis for analysis of a "behavior problem," which often has its origin in complex interactions among people and environments (see Cantrell & Cantrell, 1975; Smith & Grimes, 1979; Prieto & Rutherford, 1977), but it is a critical component of the total assessment.

The purpose of this paper is to introduce several methods of systematic observation and to comment on their relative advantages and disadvantages in order to assist the beginning observer to make a thoughtful choice of the procedure that best meets his/her needs. Examples are provided. Fuller discussions of the various theoretical perspectives on observation are available in Weinberg and Wood (1976). Boehm and Weinberg (1977) have written a brief, nontechnical manual that expands on several of the topics mentioned here; they also discuss such issues as reliability and validity which relate to observation as measurement. For a discussion of applications of observation to research, see Hersen and Barlow's (1976) excellent book.

Some Key Ideas

Before proceeding to the discussion of methods of observation, several key terms which will be used throughout this paper need explanation. A good grasp of their meaning will help the reader better understand the discussion. These terms are observation, observation records, description and inference, valuing, and labeling. The reader will note that they have to do not only with the act of observing, but with what one does with those observations.

Observation is of course the key term. What is good observation? How does observation differ from locking? All humans spend much of their waking hours looking at things in the world around them; all people are, in some sense, observers. The difference between "lookers" and "observers" is system. Good observers have trained themselves to watch carefully for easily overlooked patterns of activity. Good observers are careful, systematic "lookers."

One can never remember everything that he/she has observed. People quickly forget the details of what they have seen and find that memories alone are often unreliable. Therefore, people who need to use observation for teaching purposes find it helpful to make observation records. Later in this module, several different kinds of observation records will be described. It is important to note early on, however, that the method chosen for recording observations also influences what is observed. By choosing a method of recording that will be sure to provide the information needed to help in decision-making, one can be more certain that the observations are directed toward what is important in the situation observed rather than being haphazard.

A good observation record is a careful, accurate description of some of the actions of persons in the situation observed. It is a description of what they do, their directly observable or "overt" actions. Their thought and feelings are of

interest to the observer as well, but thoughts and feelings cannot be observed directly. Because thoughts and feelings are "covert," what is said about them results from inferences, which are based on observations of actual behavior, but which must not be confused with the observable behavior itself. In everyday activities, one is continually observing, inferring, acting, observing effects, predicting outcomes. It is often difficult to separate actual observations from the cognitive operations that are performed using them. When one uses information obtained through observation for educational decision-making, however, it is important that sincere effort be made to separate observations from what is done with them, especially in regard to the drawing of inferences. If one cannot separate them, he should at least learn to recognize when he is describing, when he is inferring, and when he is combining both processes.

People never observe without placing a value judgment on what is observed, or valuing it in a negative or positive sense. Students are observed in school settings because someone considers their behavior of concern. Valuing enters into observations of students with school-related problems very frequently. But here again, one needs to try to keep the valuing of the observations separate from the descriptive notes in the observation records themselves. One needs to learn to describe behavior first. Stating that it is "good" or "bad" should come later. Unfortunately, this is not the typical case. Usually, one begins with a negative evaluation of a student's behavior based on haphazard observation. Thus, someone reports that John fights with other students. Since fighting with other students is generally considered "bad," the observer sets out to observe carefully how often John fights and to keep a record of the number of times. Thus, the value judgment has shaped the choice of what kind of record to keep, and that in turn will direct the observations.

Valuing of observed behavior leads to the application of "labels." Labels can be useful symbolic shorthand. Labels are dangerous, however, when they are used carelessly, forgetting how much observable behavioral detail is subsumed under a single word like "emotionally disturbed" or "behaviorally disordered." Labels such as these have a powerful impact on the educational experience of the students to whom they are applied, and are often used carelessly. One must be on guard about this, because when someone is given a label such as "emotionally disturbed," it is too easy to think one knows what he/she is like. "Oh yes, emotionally disturbed. What a problem!" The more detailed and specific observational records of students are, the more will be known about their actual behavior and the less will be the need to rely on these broad labels.

ANECDOTAL RECORD OR CRITICAL INCIDENT TECHNIQUE

Most educators have recorded observations of "critical incidents" in their classroom in a narrative, anecdotal record format. A discussion of this procedure will introduce several important principles of good observation technique. Using the anectodal record procedure teachers describe in narrative form incidents of a student's behavior which are considered especially noteworthy. Most such records are of disapproved behavior; the narrative illustrates the student's particular problem. In the following narrative, the behavior of a student referred by his teacher as distractible and hyperactive is described:

> During today's math seat work time at 10:30 a.m., Bill was restless and distractible as usual. He kept turning around in his seat and bothering his neighbors. I had to speak to him several times, but he always started up again in a few minutes. Eventually, he began talking to some other children. When Mary refused to pay attention to him, he poked her with his pencil. By now his behavior had become so out of control I had to ask him to leave the room and sit on the bench outside Mr. Brown's office. This same sort of thing goes on every day. He just won't sit still or pay attention!

As the teacher's indictment of a student whose behavior is a problem for her and some of his classmates, this statement is useful. As a systematic description of his behavior, it is poor. As an assessment providing a basis for planning ways to ameliorate the situation, it is of little use.

The account begins well with the record of the time of day and the activity. However, because like most critical incidents this one was written at some later time, the time of day is only approximate. The events described are not related to specific times of the period.

PRINCIPLE 1: Observation data should be recorded at the time they are observed.

PRINCIPLE 2: Observation data should be continuously related to the time the observed events occur.

If one is to obtain an unbiased description of a student's behavior, then it must be planned in advance when the student will be observed, and observations must be recorded continuously during that period, linking them to the passage of time. If one notes down only the times when the student's problem behavior is at its peak, the record will tend to be biased toward merely documenting of the observer's pre-existing negative evaluation of the student's behavior.

The preceding sample record fails to separate inference from observed behavior -- a very common problem. Inferences are often value laden. For example, Bill is described as "restless and distractible as usual" and "out of control." Despite the strong tendency to agree immediately with such descriptions, one must take a closer look at them. On what behavior are these statements based? Instead of a description of how Bill behaves one has the teacher's evaluations of that behavior. Furthermore, there is no basis for comparing Bill's behavior with that of his peers. Are they also "restless and distractible"? How much more "restless and distractible" is Bill than they are? The issue of normative behavioral observation is also discussed in the attached readings.

PRINCIPLE 3: An observation record should be based on observable behavior. Inferences about thoughts and feelings, which are not directly observable, and evaluative comments should be linked to a description of the behavior on which they are based.

PRINCIPLE 4: Quantitative statements about whether an individual student behaves in a particular pattern more or less than peers should be referenced to observations of peer behavior made at the same time and place.

Would everyone have agreed with Bill's teacher that he is restless and distractible? The question is hard to answer. What does the teacher mean by "restless and distractible"? Would not it help if there were some agreement in advance on what those terms meant?

PRINCIPLE 5: Descriptions of behavior should be phrased in terms that are well defined. These definitions should be communicated along with the description of the behavior.

PRINCIPLE 6: Whenever possible, more than one observer should describe the same behavior. If this is not feasible, observers should practice together or test themselves against an expert standard to avoid the inadvertent intrusion of idiosyncratic bias into their observation records.

NARRATIVE RECORDS AND LOGS

Principles 4, 5, and 6 are not easily accomplished when the anecdotal record technique is used. But, this procedure for recording observations can be a more useful source of information than the first example if Principles 1, 2, and 3 are applied. Such a record will be called a narrative record to differentiate it from the post hoc anecdotal record. Below, the reader can see how the sample incident might read if the recording were made as the incident unfolded, the time intervals were noted more frequently, and evaluative terminology were avoided.

10:30 a.m.: Class doing math seatwork.

10:33: Noted Bill looking up from work, drumming with pencil on desk.

10:35: Bill turns 90% from front and leans toward Mary's desk. He turns back and looks down at work when I look at him and frown.

10:38: (Had been busy helping George at front of room for several minutes.) Bill has turned toward Mary again and is saying something I cannot hear to her while he looks at her paper. Mary tries to turn her back on him and moves her paper over to the other side of her desk. Bill pokes her back with his pencil. Mary slaps back at Bill who yells out that she is a "b____" and he doesn't want to see her paper anyhow. He is standing up now, and as I move toward him, he backs down the aisle, kicking at desks and chairs as he goes. I point toward the door and he opens it and backs away down the hall toward Mr. Brown's office. I tell him to sit on the bench and go in to explain to Mr. Brown why he is there. When I come out again Bill is sitting on the bench but will not look at me or speak to me. I tell him he is to stay there until I come to get him. Bill turns away and says something I cannot hear.

10:43: Back in class. There is some noise and talking going on as I come into the room but the students turn back to their work when I come in.

This time, the teacher has given an account of direct observations of Bill's behavior rather than a statement primarily describing reactions to that behavior. It is possible for teachers to write such a record if they are trained (or train themselves) to observe, make brief notes at the time, and write out a complete description of the incident as soon as possible. Used by well-trained, experienced observers, who often employ special techniques to facilitate the recording of as much detail as possible, the narrative record rivals the videotape in its approach to the unattainable goal of a complete record of behavior and has the advantage of flexibility of use in field situations

where accurate sound recording is often impossible. It is difficult to write narrative records without resorting occasionally to the use of inference, however, as this segment from a narrative record written by a skilled observer illustrates:

> (The subject is Patrick Taylor, a fourth-grader, who walked on crutches with his right foot carried in a sling, due to Perthes disease.)
>
> Time: 10:22 a.m.
>
> Noticing this, Patrick jumps up, since he's the next one to bat, and swings on his crutches over toward the backstop.
>
> As he approaches, Glen walks toward him and says in a friendly, cheerful way, "I'll run for you."
>
> At the same time, Harry calls over from the bank, "Hey, Patrick, can I run for you?"
>
> Patrick turns around and, without responding directly to either of these two boys, picks up a bat from the ground.
>
> He looks at Ken, who's standing nearby, and speaks to him in a quiet voice. I can't hear what he says.
>
> He obviously asked Ken if he'd run for him, however, for Ken runs immediately over and takes his running stance just a few feet away from home base on the first base line.
>
> Patrick hops with the bat directly up to the plate, leaving his crutches by the backstop.
>
> He stands there, balancing deftly on his left foot, with the bat perched eagerly up on his shoulder, ready to bat.
>
> Just before he reaches the plate, the pitcher calls in to him with slight impatience at the delay, "All right, let's go Patrick!"
>
> The first pitch comes across, and Patrick swings. The bat connects, but it's a high foul ball which the catcher cannot get his hands on.
>
> Patrick watches the ball's trajectory, then hops around, swinging the bat for practice as he waits for the pitch.
>
> The ball goes back out, comes in again, and again Patrick connects. This hit goes down toward the first base but outside the line for another foul.
>
> The runner, not being sure, goes aprt way down toward first base to play safe.

Note: From "An Ecological Study of Children with Physical Disabilities in School and Home" by P. Schoggen. In R. Weinberg & F. Wood (Eds.), Observation of pupils and teachers in mainstream and special education settings. Reston, VA: Council for Exceptional Children, 1975.

For more ideas about narrative recording see Schoggen (1964) and the material in the "Activity 1" intended to provide students with experience in doing narrative descriptions of the behavior of an observed student.

Narrative recording, particularly of the anecdotal type, is the method of observation recording used most frequently by teachers, probably because it seems a "natural" way to record, one that requires little special training or practice. Teachers often keep running anecdotal records on students about whose behavior they are concerned or narrative logs. Actually, as the example illustrates, the best narrative recording requires training and skill and is a full-time task for the observer. There are other methods of recording that are more easily used by teachers, and while none of them meet all of the principles of good observation practice listed above, the records obtained will be of better quality than the typical informal anecdotal record.

BEHAVIOR, TIME INTERVAL, AND FREQUENCY

To make observation recording easier, one must sacrifice some completeness in the record and accept the necessity of making inferential judgments as one observes. Consideration of certain key variables will assist the observer in making good decisions about the choice of an observation method. The key variables needing consideration in planning an observation system are the behavior patterns to be observed, the time interval in which they are to be observed, and the frequency of their occurrence. When observing some behavior patterns, it may also be important to observe their duration, that is to say, how long a given occurrence lasts, but the specification of which, when, and how often will cover most cases with which teachers are concerned.

Before turning to a more detailed study of some examples of procedures developed for observing and recording behavior, the reader is referred to Figure 1 which summarizes the options, showing how one can vary the difficulty of the observer's task and the ease of recording along the dimensions of inclusiveness and detail of description.

	Complete description		Partial description	
More difficult to observe/ record	Short time intervals (continuous)	↔	Long time intervals (time sampling)	Less Difficult to observe/ record
	Many behavior categories (exhaustive)		Few behavior categories (focused)	
	Frequent occurrences		Rare occurrences	

Figure 1. Dimensions of inclusiveness and detail of description in the observation and recording of behaviors.

In the following discussion, the author reviews some procedures that vary along these dimensions, beginning with those less difficult for the teacher/observer to use and proceeding to some of greater difficulty which have features that may recommend them for special uses. Each will have its advantages and disadvantages over anecdotal and narrative records.

A Very Simple Record

One may begin with a very simple system. Since writing out a narrative description of behavior takes all the observer's time, i.e., the observer attempts to record as much as he/she can without limiting in any way the kinds or frequency of behavior to be recorded, one may decide in advance that all social behavior observed will be classified simply as either "acting-out" or "withdrawing." Most problem behaviors seem to cluster under the one or the other of these two categories. They are rather general terms, and to apply them one must use some inference, but they can be pinned down by stated, specific examples. For example, "acting-out"

includes such behavior as hitting people or objects, yelling, making loud noises, and the like, whereas "withdrawing" includes turning away from others, not speaking when spoken to, and the like. If the observer decides that he/she will check a child only once a day for either kind of behavior, the time interval is one day. Thus, the observation record might consist of a sheet divided into two columns, one headed "acting-out behavior" and the other, "withdrawing behavior," in which is entered a tally mark each day on which the student's behavior is observed. The total of the tallies would be the frequency of occurrence of the behavior. Going back to the narrative description example again, it looks as if one would have placed a tally mark in each column to record observations of Bill's behavior on that particular day. He acted out at sometimes, but he withdrew at others.

This system is limited, of course, by its very simplicity, but it has some advantages over the use of anecdotal records and is less time consuming than a narrative recording procedure. For example, the behavior of selected peers can be easily observed and recorded on the same basis so as to provide a context for considering the relative uniqueness of Bill's behavior. Also, one can check his/her observations against those of other observers, thereby obtaining an estimate of interrater reliability.

The same advantages are retained when one begins to refine the system by breaking down general behaviors into more specific behaviors and/or longer time intervals into shorter ones. Since the number of behaviors x the number of intervals = total possible frequency of recording each day, the possible frequency totals rapidly increase. In the example above, the total possible was only two. But, suppose a teacher decides to record the occurrence or nonoccurrence of three behaviors each morning and again each afternoon, or twice a day. The total number of tallies possible now increases to six, a small increase to be sure, but the trend is clear. Even simple yes/no decisions about the occurrence/nonoccurrence of an event take thinking and recording time. How many such yes/no decisions on

the occurrence of a particular behavior can a teacher manage each day? 12? 100? 500? At what point is the increasing number of tallies no longer a useful addition to the record and just more work?

There is no definitive answer to this question. Each problem one seeks to solve requires the reconsideration of the trade-off between number of kinds of behavior to be recorded and length of the observation interval. The specifics of the system should reflect the purpose for assessing each individual student. As a practical matter, useful observation records range from 1-200 possible tallies per day for a given individual. Reviewed below are several specific observation systems that will serve as examples of the further application of the ideas just discussed.

Behavior checklists

Behavior checklists or rating scales typically list 50-60 commonly observed behavior patterns. Usually completed on a time interval varying from a week to a year, the observer checks those behavior patterns that characterize the person observed. Most scales provide for a general estimate of frequency of occurrence by providing space to indicate whether the checked behavior occurs "never," "seldom," or "often." One might say that in such a scale the two broad categories of behavior, "acting out" and "withdrawal," have been broken down into many separate behavior patterns, thus making more work for the observer but providing a more complete record. Examples of well-known, commercially available rating scales are the Quay-Peterson Behavior Problem Checklist (1979), the Walker Problem Behavior Identification Checklist (1976), and the Devereaux Elementary School Behavior Scale (Spivack, & Swift, 1967).

In completing a rating scale, an observer typically makes numerous retrospective inferential judgments about the relative severity and duration of

the behavior observed. Below are some sample items from the School Behavior Profile, a representative checklist developed by Balow and Rubin (1974) as part of a longitudinal study of the social and academic behavior of students.

Nervousness, jitteriness, jumpiness. (inference required.)

Hyperactivity, hardly ever still. (inference required.)

Doesn't speak; uses only grunts or noises to communicate. (descriptive, inference required only in estimate of frequency.)

Anxiety, general fearfulness. (inference required.)

Steals. (descriptive, inference required only for estimate of frequency.)

Disruptiveness; tendency to annoy and bother others. (inference required.)

The lack of a full definition of the descriptors used and the absence of provision for recording a description of the specific observed behavior on which the rating is based tends to decrease the usefulness of such checklists as observational records. Checklists are also lacking in focus on specific behaviors of concern in the individual case. Observation checklists and scales are probably most useful for collecting general impressions of behavior from parents or teachers who have not been trained to provide more detailed descriptions of behavior. Such ratings should not serve as the basis for decision-making about students. However, they can be made more useful if the list of items to be checked is more carefully tailored to fit the needs of a particular group of students or a particular program. Such a list may also be shortened or simplified by dropping descriptors not relevant to the problem behavior of any particular student or group of students with whom one is concerned. Responses to items that require considerable inference should be viewed cautiously. "Depression," "anger," and "anxiety" are examples of internal states that can be revealed through many different overt behaviors and masked by many others.

Simplified Event Recording

Behavior checklists and rating scales make observation recording easier by setting a limit on the number of behavior patterns to be described, making it easier to record frequency of occurrence/nonoccurrence, and using a long time interval. As suggested, as the number of behaviors to be checked is reduced, the task of recording becomes easier.

There are other ways to simplify the observer's task and save time for recording other information about the behavior observed. It has been noted that the actual frequency of occurrence of teacher-disapproved behavior is usually much lower than the incidence of approved behavior. Since the necessity for making a tally mark adds appreciably to the time required to make a yes/no decision, some observers record only the frequency of occurrence of disapproved behavior. For example, since students spend more time working ("on task") than not working ("off task"), tallies are made for instances of "off task" behavior; since more time is spent "not hitting" than "hitting", a record is made of the instances of hitting. One problem with such a system is that it focuses the observer's attention on the student as a producer of disapproved behavior and may eclipse awareness of his more typical "good" behavior, but as one soon learns about classroom observation, "to get one gives". No system is without faults which must be taken into account by the thoughtful observer.

By eliminating behavior descriptors according to one or more of these rationales, experienced observers have developed useful general systems that require the recording of the frequency of only a small number of kinds of behavior. Forced by circumstances to choose only one behavior to be observed, and defining it negatively to minimize the number of tallies to be recorded, a good choice would be the critical school social behavior of being "off task," that is, not attending to one's work. "Off task" is defined as not looking at the assigned task, but

obviously subsumes many social behaviors that are inappropriate when work is to be done. Deno and Mirkin (1977) suggest four critical social behaviors to observe: "off task," "noise," "out of place," and "physical contact." This simplified event recording system will be described in more detail later in this module.

Simplifying by Varying the Time Interval for Recording

Besides varying the number of behaviors observed to change the difficulty of the observer's task, the length of time intervals for recording can be adjusted to fit the observer's available time. As already noted, behavior checklists are commonly used to summarize observations made over long intervals; but they can also be used to cover intervals of any length. In practice, however, one finds that as the intervals become as short as only a few seconds, a point is reached at which only one behavior pattern can occur within an interval. By using intervals of 2-3 seconds, then, one gets an almost complete continuous record of the frequency or rate of occurrence of an observed behavior. Only highly skilled observers can maintain such a record on more than a few behavior patterns. Usually, however, the possibility of such a complete record is given up in order to make the task easier, and the observer simply accepts the resulting loss of some information. How critical this loss may be can only be answered in regard to a specific problem.

Intervals of any length can be used. If data collection will continue for weeks or months, entering one tally for an interval of as long as an entire daily class period is adequate for a record of some behavior patterns such as attendance at school or coming to class prepared. As a general rule, intervals of from 20 seconds to 1 minute permit observers to record accurately the occurrence of several key behaviors and meet most needs adequately.

As mentioned previously, a behavior checklist may be regarded as an extensive list of behavior patterns which is completed by an observer as a

summary record at the end of a relatively long period of observation. However, such checklists could be completed more frequently, daily or hourly, but since they are extensive and include many patterns that are not relevant in an individual case, such frequent completion would be awkward and time consuming. Other than paring the list down to a handful of behavior patterns to fit each individual, is there some way to make it more manageable? The following strategy has been used successfully:

(a) Develop a list of social and academic behavior patterns that are considered important within a school setting and/or to the development of a particular student or group of students; (b) Arrange the items in an order based on a typical sequence of human development or a logical task hierarchy (in either case the list will move from basic or simple behavior patterns to more complex patterns that incorporate additional elements); (c) Start at the beginning of the list and check off the behavior patterns observed performed by the students being rated, eventually reaching a point on the list where the listed behaviors are not being demonstrated by a particular student; (d) Choose the first five or six missing behavior patterns as objectives for instruction and maintain a daily, individualized summary record of student performance, only periodically reevaluating students against the entire list.

This procedure has the work-reducing advantages of a relatively long time interval (school day or class-period) and relatively few behavior categories to be attended to at any one time while it retains the potential of an exhaustive description of desirable behavior shown by the student observed. Just such a procedure was used by M. M. Wood and her colleagues at the Rutland Center, Athens, Georgia, to develop the Developmental Therapy Objective Rating Form (DTORF) for seriously handicapped students (1975, 1979).

In using such a procedure, teachers and their assistants move down a list of behavioral objectives during the initial assessment, checking those behavior patterns they have observed performed appropriately by the student 90% of the time. Instruction is then directed toward student mastery of a target group of six behavior patterns in each category. Each afternoon, after the students have gone home, the teacher/assistant team discusses the students' behavior and makes a daily record of whether a specific behavioral objective was performed "always," "sometimes," or "never" by a given student. When mastery is shown for a sufficient period of time, the mastered behavior is checked off the student's list and attention turns to the next missing behavior pattern. Thus, teachers observe and record the occurrence/nonoccurrence of approximately 30 behavior patterns for each student during each one-day interval. Summary observation information recorded using this procedure provides an important data base for the program, both guiding instruction and permitting evaluation of program effectiveness.

Similar procedures using lists of behavioral descriptions/objectives have been described by Hewett and Taylor (1980) and Stephens (1975). Such lists are often used with intervals longer than a single day; however, the procedure described seems best suited to providing an observation record immediately useful to teachers of students with special needs. In developing such a list of behavioral objectives for mainstreamed students, regular class teacher's should feel free to consult special education personnel. In most instances the best time and place for developing such a list is at the IEP conference where regular and special eduction teachers, school psychological staff, administration, and a student's parents meet to formulate goals and objectives for the student and a plan to realize these goals and objectives.

OBSERVATION SCHEDULES: DENO AND MIRKIN'S FOUR BEHAVIORS

As already mentioned, Deno and Mirkin (1977) illustrated another approach to developing a procedure that provides useful information to the teacher without requiring excessive time. Based on their experiences as consultants to class teachers working with special education students, they selected a group of only four social behavior patterns that "fairly represent the 'categories of concern' for most classroom teachers" (p. 101). As already mentioned, Deno and Mirkin defined the four patterns negatively to make the recording of frequency of occurrence less difficult: "noise," "out of place," "physical contact," and "off task." They suggested that the target student be observed initially 10-30 minutes each day for 5-7 days and, thereafter, as frequently as possible. The four categories are defined below (From Deno & Mirkin, 1977, pp. 101-102).

1. Noise: Any sounds created by the child which distract either another student or the teacher from the business at hand. The noise may be generated vocally (including "talk outs" or unintelligible sounds) or nonvocally ("tapping a pencil" or "snapping fingers").

2. Out of place: Any movement beyond the either explicitly or implicitly defined boundaries in which the child is allowed movement. If the child is seated at his desk, then movement of any sort out of the seat is "out of place."

3. Physical contact or destruction: Any contact with another person or another person's property which is unacceptable to that person. Kicking, hitting, pushing, tearing, breaking, taking, are categorized as physical contact or destruction.

4. Off task: Any movement off a prescribed activity which does not fall into one of the three previously defined categories. "Looking around," "staring into space," "doodling," or any observable movement off the task at hand is included.

Some teachers may find they can make such a record while teaching or carrying out other academic tasks. Most will find they will need the assistance of an aide, the principal, or a school psychologist. Deno and Mirkin suggested that this procedure be used to help spot critical areas of problem social behavior. More specific pinpointing of behavior patterns for change would then follow.

Deno and Mirkin also suggested recording behavior continuously, that is, noting each occurrence of one form of the behaviors. This is called frequency or event recording. For accurate recording, events must have a class beginning and ending point. Not all behavior of interest can be easily divided into such units. Anyone wishing to make the use of this system less difficult could specify observation intervals of one minute or 30 seconds. The record would then consist of the number of intervals in which the behavior pattern was observed to occur (Fig. 2), a number that usually will be less than the actual frequency of occurrences.

	1st min.	2nd min.	3rd min.	4th min.	5th min.	Total
Frequency record	//	/	///	---	/	7
Interval record	/	/	/	---	/	4

Figure 2. Occurrences of "off task" behavior.

The observer's task is easier when exact frequencies are not tallied, but some information is lost. Neither of these procedures provides an accurate measure of the duration of behaviors, which may sometimes be of interest. For each application, observers must ask themselves how much detail is useful. There is no reason to collect more observational data than can be or will be used.

A LONGER PUPIL OBSERVATION SCHEDULE

Previously in this module, the writer stressed the desirability of limiting the actual observation record to observable behavior while leaving inferential and valuing statements to a discussion section of the record or adding them as commentary. The critical reader may have noted that evaluative terms tend to creep into checklists and behavior pattern descriptions such as those mentioned. "Noise," in the Deno/Mirkin category system, is "sounds which distract ...." (p. 101). "Physical contact or destruction" is "contact which is unacceptable (p. 102)." Interpretation of the terms underlined requires a value judgment by the observer.

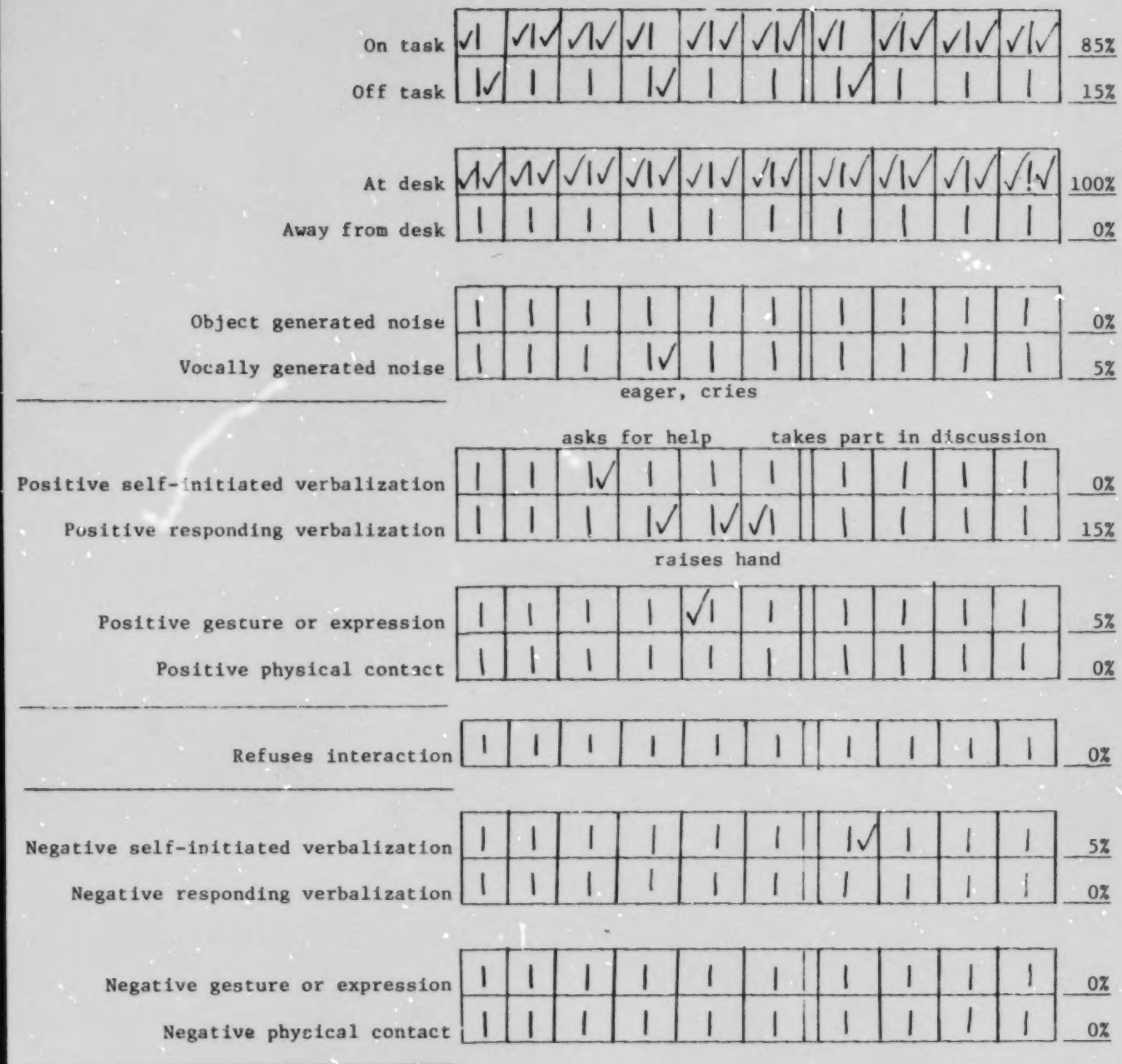
Records of the incidence of behavior patterns to which such inferential judgments have been applied always should be somewhat suspect because any inference reduces interobserver reliability; that is, people usually will agree about the times they hear a student speak to a classmate but sometimes will disagree about whether the verbalization was "positive" or "negative" in intent.

The Pupil Observation Schedule (POS), a procedure developed by the author (Wood, 1973, 1979), has several categories that require observers to infer whether a student's or teacher's intent is positive or negative. The POS permits recording the occurrence of 13 categories of student behavior by one-half-minute intervals. The observer makes a check mark in the cell for each half-minute

whenever a behavior pattern occurs one or more times during that interval. This method results in some loss of data, as illustrated previously, but after several observation sessions a satisfactorily complete picture of student behavior can be obtained. With this type of system, it is customary to summarize the observations as percentages (number of intervals in which a behavior was observed to occur/total number of intervals in the observation session). The percentages have been calculated on the illustrative POS form (see Figure 3) which shows the record for a 10-minute observation session (20 30-second intervals).

The first four categories of behavior recorded on the POS can be objectively defined: "on task," "at place," "object noise," and "vocal noise." ("Off task" is included so that the observer always will have to mark a student as "on" or "off task," thus marking off an interval even if no other activity is observed.) "Non-response" to a behavior initiated by another is also a category that requires little inference from an observer. But "self-initiated verbalization," "responding verbalization" "gesture," and "physical contact" must be judged either "positive" or "negative" in intent. Some objectivity of recording has been sacrified in the interest of reducing the number of categories.

Observers can learn to use the POS with only a little practice. Because of the long time intervals, observers usually find it possible to elaborate on the record with brief narrative comments as shown in the example. The procedure



Time: 10:15-10:25 a.m.

Date: 12/3/78

Activity: Arithmetic period. Teacher is demonstrating problem solutions at chalk board. Students are volunteering answers and responding to teacher questions.

Room: 203 - Grade 5 - Mr. B.

S: C.N. - Male - Age 11-4

Figure 3: Pupil Observation Schedule (Wood, 1973).

requires the observer to be free of other responsibilities, however. It has been used successfully by resource or consulting teachers as part of their assessments of problem behavior in regular classrooms (Rardin, 1976). Rardin supplemented the pupil categories by adding five categories of teacher behavior: "control," "organize," "discuss," "demonstrate," and "describe." She also provided space for recording if the teacher attends positively, negatively, or not at all to the target student during each interval. Her modifications illustrate the desirability of fitting any procedure to the needs of the observer.

Teachers seem to find the information recorded on the POS and similar schedules useful in helping them to think about strategies for managing problem behavior. It provides a more complete and continuous record of observed behavior than lists of behavior descriptions/objectives, but the record lacks the obvious implications for instruction that behavioral objective checklists provide. The purpose of an observation schedule such as the POS is to give a general picture of the student's functioning preliminary to focusing in on specific behavior patterns of interest or when monitoring progress over time. The POS record covers a wider range of behavior than the Deno/Mirkin "four category" procedure and includes a record of positive as well as negative behavior. Walker (1979) describes application of similar observation systems to pupils acting out behavior. However, it requires somewhat more skill from the observer and cannot be used by a teacher who is simultaneously carrying on instruction. By now, it should be clear that "observation" can be, in fact must be, shaped to fit one's needs, skill level, and time schedule. Activity 2, attached at the end of this section of the module provides students with experience in using the Pupil Observation Scale. After that experience it may be useful to discuss with students these issues.

ADDITIONAL DATA THAT MAKES THE OBSERVATION RECORD MORE USEFUL

When used by a practiced observer, the narrative record procedure produces a very complete record of behavior in context which is relatively unbiased by preconceived categories. Description can be kept well separated from inference most of the time, although as illustrated, it would be false to claim that subtle observer biases do not influence what goes into the record. However, the full narrative procedure is difficult and time-consuming to use, and a narrative record is not always the most useful kind of observation record for the teacher since the data recorded is difficult to summarize and apply to the solution of instructional problems. One problem facing observers, then, is how they can restore to observation records as much as possible of the context captured in a complete narrative record.

One constructive step is to prepare a brief narrative description of the classroom environment, the activity being carried on, and the characteristics of the students and teachers (i.e., age, sex, socioeconomic status, and skill levels). In addition, observation records should be made of the behaviors of a random selection of the target student's same-sex peers. A common procedure is to alternate observations of the target student and individual peers; that is, observe the target student for 5 minutes, peer one for 5 minutes, the target student again for 5 minutes, peer two for 5 minutes, and so forth. Studying such a record helps one judge the "normality" of the target student's behavior. Forness (1979, attached in this module) describes ways of collecting and using such normative data. As already mentioned, Rardin found it helpful to include a partial record of the teacher's behavior on the POS form, thus adding another piece to the description of the context in which the student's behavior occurs.

Although the Pupil Observation Scale provides a way to start recording observations of social interaction in the classroom, the other observation recording procedures described in this module have tended to focus on the individual student. Recording interaction is more difficult but, as Strain, Cooke and

Appolonni (1976) have pointed out, interaction data adds much to an observation record. As an observer gains in skill, he/she finds ways to add important notes on social interaction to the record by annotation or modification of the form being used.

Below is another simple procedure to record interactions between small groups of individuals. An observer can draw a circle containing the names of each student being observed and draw arrows to show the direction of an interaction, as shown in Figure 4. Keep a running tally of positive (+) or (-) interactions during the period of observation.

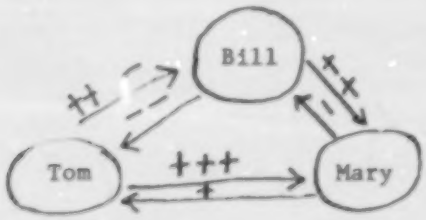


Figure 4. A method of observing social interactions in classroom settings.

The same procedures that have been used to record the behavior of individuals can be used to record group behavior. A group record is simply the sum of all the individual behaviors of a particular kind, for example, "verbal offers to help," or "physical assistance". Such group summary records can be used to plan interventions to change group behavior and monitor their effectiveness or to place individual behavior in context. Usually the total number of behavior patterns to be observed and recorded is kept small or the record is made only at a specified time interval, for example, a record is kept of the number of students "off task" at each quarter hour.

Reliability and Validity of Observation Data

The easiest way to shake the confidence of the novice observer to whose needs this paper has been addressed, is to ask him or her "to demonstrate the reliability and validity of their observation record." Records should be as reliable and valid as possible, but like so much else, the necessary degree of reliability and validity varies with the particular application.

The reliability of a record concerns the accuracy with which it has been made. Only subjective estimates can be made of the reliability of narrative records. Reliability of count data (event or interval) is usually reported as the percentage of observer agreement. The easiest way to calculate this percentage from event data is to divide the smaller number of occurrences recorded by the larger number of events recorded by two observers during the same observation period, then multiplying by 100 to convert the ratio obtained to a percentage. Example: Observer A records a student as "off task" eight times during a ten minute observation period, while Observer B records the student "off task" only six times. Their interobserver agreement is only 75%, which is not particularly good reliability. An agreement of at least 80% is generally considered desirable.

If several behaviors are being observed simultaneously, the percentage of agreement can be determined for each behavior. These percentages can be averaged to give an estimate of reliability overall. Such a percentage gives equal weight to all of the individual scores, and since reliability of observation and frequency of occurrence will vary for each behavior, a weighted average may be desirable. Boehm and Weinberg (1977) describe a simple procedure for calculating such a weighted average. Their procedure is also useful for determining the percentage of agreement when more than two observers are being compared.

A better way to calculate interobserver agreement for interval data is to divide the total number of intervals less those where there is disagreement by the total number of intervals, then converting this ratio to a percentage. Example: Observer A and Observer B agree that the student they observed was "off task" during eight of ten intervals, but Observer A marked one interval B did not and B marked one interval A did not mark. The percentage of agreement in this case is 8/10 or 80%.

If the percentage of agreement is less than 80%, one may suspect that the events to be recorded were poorly defined, so that different observers interpreted their occurrence differently, that too much inference was required in deciding whether or not a particular event occurred, or that the mechanics of the recording procedure were so complex as to cause observer recording errors. If the system does not have major flaws of these kinds, reliability can be improved by practice.

Reliability places a limit on the validity of observation records. Validity is the truthfulness of the record. Does the record describe what actually happened? If the description is inaccurate, it is clearly not valid. But validity is also a affected by other factors. For example, a valid record would have to be a complete record. In this module, it has been suggested that observers should only collect data that is useful. This practical advice clearly places a limitation on the validity of the records obtained. Inference also affects validity. For example, even though two observers may agree that a student's behavior reveals "depression" or "anxiety," a third observer may feel that the same behavior should be interpreted as "anger." While this may show up as a reliability problem, the fundamental problem is one of validity. Thus, an objective description tends to be more valid. Yet, to make use of the data, someone must interpret and value it.

Obviously observers must be concerned about the reliability and validity of their records, but not to the point of giving up the simple methods of observation recording that have been described. Despite their technical flaws, the chances are great that records made using these methods are more reliable and valid than the haphazard oral or written anecdotal records that too often provide the foundation for decision making and evaluation of individual and group educational programs.

USING OBSERVATION RECORDS

A variety of procedures to guide observation and make a record of what is observed have been described in this paper. How does one decide which procedure is most appropriate for a specific purpose? Some teachers may prefer to adapt the same basic procedure to meet every need, but in general, it is better to master several different approaches so as to have a range of choices when seeking for the "best fit." Perhaps, members of a school faculty can pool their expertise, some becoming experienced with one recording procedure and others with another. Here are some possibilities.

1. Using observation for screening and initial assessment: Observation is an important tool for use in answering the first question teachers should consider when they first notice that they are "disturbed" or concerned about a student's behavior: How specifically can they describe the behavior that is disturbing them? Careful description must precede decisions about when and how to intervene, or when and to whom to refer. Brief narrative notes taken after the occurrence of "critical incidents" can be of help in pinpointing the problem, but early use of systematic observations such as the Deno/Mirkin "four category" procedure, or a form of the Pupil Observation Schedule will provide data on the "typical" as well as the "problem" behavior. Both should be considered. Use of these procedures can and probably should be relatively informal at this stage.

After the teacher begins to feel more certain about what specifically is the behavior that disturbs him/her, the next important question can be addressed: Is the behavior sufficiently disturbing to the student, the group, or to the teacher to warrant planful attempts to change that behavior? Most problems will respond to good classroom management procedures; many are transient and seem to be "outgrown" without special attention. Is one reasonably certain this is one of the small number that requires a special plan? If so, procedures described in the modules in this series dealing with topics such as referral, assessment, development of individual and group behavior management procedures, and the development of behavioral objectives and Individualized Educational Programs (IEPs) should be undertaken.

2. Formal assessment and the development of the Individualized Education Plan: The systematic observation procedures that have been described under Step 1 are appropriate for use in formal assessment and program planning as well. But, for these purposes, procedures like Wood's Developmental Therapy Objectives Rating Form (DTORF, 1975) and Hewett and Taylor's ABCs of the IEP (1980) become particularly useful. Too often educators wait to use these procedures only after the student has been placed in a special setting. Since the relationship between changes in settings and changes in behavior is well known, it would seem appropriate to rate a student's behavior on such scales in the situation where the problem was first observed as well as later, after attempts have been made to change his/her behavior or after the student has been placed in new or different situations. Such a "double baseline," if we may so term it, provides both a helpful summary of the student's original status and appropriate objectives for the early stages of a behavior change program when a detailed plan has not been worked out.

3. Using observation procedures to monitor progress: Several methods for obtaining frequent independent observations of pupil behavior have been described

earlier. Data from such observations can be readily summarized on charts or graphs (Deno & Mirkin, 1977). Thus summarized, the data permit ongoing evaluation of the interventions introduced to guide the social development of a target student or students and to provide the basis for thoughtful program decisions. Such use is described by Deno and Mirkin (1977) as "data-based program modification."

4. Using observation procedures to determine what to change when progress does not occur: Behavior change programs are usually focused on a small range of student behavior, for example, being on task or responding appropriately to social greetings by others. By thus narrowing the focus, there is a tendency to lose sight of the "total person." While one response to a student's lack of progress may be to change the intervention plan while continuing to focus on the target behaviors, another might be to take a step back, and look at the student's behavior more generally, as was done in the earliest stages of assessment. Focusing on targets or behavioral objectives, such as those on the DTORF, may be an impediment to noticing that an individual student might respond better if approached in a completely different manner. Going back to the use of a general observation procedure like the POS may help to reassess the situation and develop new plans that work better than the old.

5. Using observation data for the purpose of evaluation: The advantages of using charts and graphs to summarize observation data have already been mentioned in the discussion of methods for monitoring progress. Such data summaries can also be useful for purposes of program or individual progress evaluation. Data from behavior sequences like those of M. Wood (1975) or Hewett and Taylor (1980) can be summarized in terms of specific objectives mastered by individual students or number of objectives mastered by a group during a specified time period. Such data can be very useful in evaluating individual and group program effectiveness.

A CONCLUDING COMMENT

Slavin (1978) argued in a recent issue of the Educational Researcher, the opinion journal of the American Educational Research Association, that special education programs have demonstrated much greater potency as interventions fostering the development of students in part because of their attention to establishing a foundation of data gathered through systematic observation on which to base program planning and decision making.

Several years of experience have convinced the writer that systematic observation is the best procedure available to the teacher who wishes to gather useful information on the social behavior of students. The procedures may seem awkward and time-consuming at first, but with practice, they become easier to use. However, the most convincing argument is made by the observation records themselves. Most teachers will find them useful enough to be worth the extra effort it may require to gather and maintain the data they contain.

References

Balow, B. E., & Rubin, R. School Behavior Profile. Unpublished rating form, 1974. (Available from B. Balow, Pattee Hall, University of Minnesota 55455).

Boehm, A. E., & Weinberg, R. A. The classroom observer: A guide for developing observation skills. New York: Teachers College Press, 1977.

Cantrell, R. P., & Cantrell, M. L. A record of ecological analysis: Heuristic case report. In W. C. Rhodes (Ed.), A study of child variance: Exercise book. Ann Arbor, MI: University of Michigan, 1975.

Deno, S. L., & Mirkin, P.K. Data based program modification: A manual. Reston, VA: Council for Exceptional Children, 1977.

Forness, S. E. Normative behavioral observation data as a standard in classroom treatment of educationally handicapped children. In R. B. Rutherford & A. G. Prieto (Eds.), Severe behavior disorders of children and youth. Monograph in Behavioral Disorders. Reston, VA: Council for Exceptional Children (CCBD), 1979.

Hersen, M., & Barlow, D. H. Single case experimental designs: Strategies for studying behavioral change. New York: Pergamon, 1976.

Hewett, F. M. & Taylor, F. D. The emotionally disturbed child in the classroom (2nd ed.). Boston: Allyn & Bacon, 1980.

Long, N., Morse, W.C., & Newman, R. G. Conflict in the classroom (4th ed.). Belmont, CA: Wadsworth, 1980.

Prieto, A. G., & Rutherford, R. B. An ecological assessment technique for behaviorally disordered and learning disabled children. Behavioral Disorders, 1977, 2, 169-175.

Quay, H. C., & Peterson, D. R. Behavior Problem Checklist. Unpublished rating form, 1979. (Available from D. R. Peterson, 39 North Fifth Avenue, Highland Park, NJ 08904).

Rardin, C. A. Comparison of the utility of two classroom observation systems as diagnostic-assessment tools. Unpublished masters colloquium paper, University of Minnesota, 1976.

Schoggen, P. Mechanical aids for making specimen records of behavior. Child Development, 1964, 35, 985-988.

Slavin, R. E. Basic vs. applied research: A response. Educational Researcher, 1978, 7, 15-17.

Smith, C. R., & Grimes, J. (Eds.). Identification of emotionally disabled pupils: Data and decision making. Des Moines, IA: Division of Special Education, Iowa Department of Public Instruction, 1979.

Spivack, G., & Swift, M. Devereux Elementary School Behavior Scale. Devon, PA: The Devereux Foundation, 1967.

Stephens, T. M. Directive teaching of children with learning and behavioral handicaps. Columbus, OH: Charles E. Merrill, 1975.

Strain, P. S., Cooke, T. P., & Apolloni, T. Teaching exceptional children: Assessing and modifying social behavior. New York: Academic Press, 1976.

Walker, H. M. Walker Problem Behavior Identification Checklist: Manual. Los Angeles: Western Psychological Services, 1968.

Walker, H. M. The acting - out child: coping with classroom disruption. Boston: Allyn and Bacon, 1979.

Weinberg, R. A., & Wood, F. H. (Eds.) Observation of pupils and teachers in mainstream and special education settings: Alternative strategies. Reston, VA: The Council for Exceptional Children, 1975.

White, O. R., & Haring, N. G. Exceptional teaching: A multi-media training package. Columbus, OH: Charles E. Merrill, 1976.

Wood, F. H. Pupil observation schedule. Unpublished observation schedule, 1973, rev. ed., 1979. (Available from F. H. Wood, Pattee Hall, University of Minnesota 55455).

Wood, M. M. (Ed.). Developmental therapy: A textbook for teachers as therapists for emotionally disturbed young children. Baltimore, MD: University Park Press, 1975.

Activity #1 - Narrative Descriptions

Narrative descriptions of the behavior of a student in a classroom situation is a useful method of assessing that student's behavior. For students being prepared to teach, practicing observation skills by writing narrative descriptions can also help develop general (as opposed to formal) observation skills. Activity 1 provides a handout for teacher education students to help them develop the "objective frame of mind" needed for writing good narratives. Also included is a sample form to help structure these observations. While videotaped or filmed sequences are useful for providing social situations to describe in narrative records, especially when used as brief training sessions, students can learn much from conducting the observation in a "live" classroom situation.

For the purposes of this practice observation, students ought to be encouraged to observe one student for a period of at least 20 minutes. The following activity outlined for use in a teacher education program has been used with both regular education and special education students at the University of Minnesota.

NARRATIVE OBSERVATION*

There are two major concerns in studying an individual's behavior. One is a factual problem: How does the person actually behave in a particular situation? The other is a theoretical problem: Why does he behave as he does? Answers to both questions are essential in arriving at an understanding. Misunderstandings arise, however, when the two processes are confused.

In general, a good narrative account of a person's behavior should be a complete and accurate account of actual behavior. A person reading the record should be able to get the same basic picture of the classroom situation as the person who actually observed it. A good narrative record should almost do what a sound film does. Writing down the behavior of one student in a classroom often seems deceptively simple. In fact it is a demanding task, especially when one later attempts to make sense out of the observations. Included below are a list of ideas that will make a narrative description better able to communicate the observed behavior of the student.

A. Things to do

1. Include time and place of observation and duration of actions.

   a. Between what times was observation conducted?
   b. Where was observation made, e.g., type of class, age of students, subject of class?
   c. Duration of observed activities. (How long did student cry, lay head on desk, do math problems? Where did student go when he/she left seat?)

2. Describe the setting in which behavior takes place.
   a. What were students expected to accomplish during observed period?
   b. How many students are in the class? How is class seating arranged?
   c. Who is leading the class? How is this being done?

3. Include relevant facts about student being observed, e.g., age, sex, noticeable physical characteristics, dress, where seated.

4. Include actual dialogue.
   (Not: Teacher sent student out of room and he refused, but: Teacher said, "Leave the room immediately and don't come back until you learn to behave." Student threw a book on the floor and said, "I didn't do it, and you can't make me go.")

5. Describe social relations involved, e.g., teacher-student, teacher aide-student, student-student.

6. Include full details of actual actions. (Not: "He drew all over his answer sheet", but: "He worked four math problems, looked around the class for 15 seconds, then folded his answer sheet over and drew the outline of a naked woman.")

*Adapted from a training module prepared by R. W. McCauley, Longfellow Treatment Center, Minneapolis, Minnesota.

B. Things to avoid

1. Confusing interpretations with behavioral facts.

   a. Some of these interpretations are inferences about motivations, such as: "He wanted to," "seeking attention"..."he was trying to"..."he was interested in"..."he feels at home in"..."he thought it was funny, so"..."he disliked." (Motives are not descriptions of behavior.)

   b. Some of these interpretations are inferences about the observer's covert behavior, such as" "he was daydreaming,""he was thinking about"..."he was unfamiliar with." (The focus in a good narrative description is on overt behavior.)

   c. Some of these interpretations are in effect theories about behavior, such as: "he has a habit of","it is in his nature to," "he has learned to," or "he has been reinforced for."

2. Labeling rather than describing.

   a. Some of these labels are "trait" names (ascribing behavior to hypothetical "traits," is but one kind of theory about behavior).

   b. Some of these labels are short-hand pseudo-descriptive attempts to summarize behavior rather than actual descriptions of behavior, such as "shy," "aggressive," "rigid," "stubborn," "inattentive," or "hyperactive." They may seem descriptive to the observer but often mean something totally different to the person who reads an account of the observation.

3. Evaluating rather than describing (related to labeling and name-calling)

   a. Some evaluations of behavior are pseudo-interpretative statements, generally loaded with the observer's feelings, such as, "spoiled," "trouble-maker," "well-behaved," "good student," "poor sport," "tyrant," "hot-headed," "mischievous," "dumb," "bright," or "disturbed."

   b. Some of these evaluations are reflections of how the observee impresses the observer rather than descriptions of what the observee does. Examples of this can be seen in observations where students are described as "likeable," "agreeable," "cooperative," "friendly," or "humorous."

4. Using words which communicate poorly

   Some terms do not describe, but only seem to describe. Examples of these are "frequently," "soon," "a short time afterwards," "large," "small," "later," and "close" or "far." (Some are actually incorrect, such as "He is <u>always</u> late," "He <u>never</u> works," etc.)

We can agree that theorizing and interpreting is necessary. We can also agree that labeling is often convenient. Evaluating others' behavior and having our feelings affected by others' actions toward us is inevitable, and at times, quite desirable. We also must understand, however, that confusions and misunderstandings results when these processes are mistaken for descriptions of behavior.

It is also well to remember that the goal of a narrative description is to write a factual account of a person's behavior, not a literary masterpiece. If the observee's behavior is monotonous, undramatic, or unspectacular, it is not the fault of the observer.

Activity #1 Narrative Observation

Instructions: Observe a child or adolescent, in some kind of school setting, for a period of twenty (20) minutes. The subject of this observation does not necessarily have to be one whom you suspect of having a school-related problem.

1. Time: Twenty-minute observation period.

2. Child: The student should be a stranger of either sex.

3. Setting: Any school situation. Acceptable places are: playground situations, classrooms, special classes such as music, gym or art, or other school activities.

4. Record: Your written report of the observations should contain:

   a. Description and factual statement of the social setting including time, place, and persons involved.

   b. Description of what is being done in the setting, including the task that the subject in the description is expected to be accomplishing.

   c. Description of the subject.

   d. A separation of observational records into four five-minute units.

5. After completing the record, write an interpretation of the meaning of the interaction between the child and his environment. What goals, desires, needs were expressed by the child or may be inferred from his/her behavior? What did he/she learn? What was he/she prevented from learning? Why did he/she behave as was observed? At this stage one can attempt to make meaning from the observational data. (As is shown on the sample observation form a column for observer's comments is included. This is a space where interpretative notes can be made as one is writing or rereading the text of the observation.)

6. Procedure: Avoid giving the child and others he is interacting with the idea that they are being watched. Make no personal contact with the child or the adults or peers.

Narrative Observation

Date/Time:

Social Setting:

Task:

Subject:

Observation Record	Observer Comments
Time: _____	
Time: _____	

(Subsequent pages need only Observation Record and Observer Comments columns.)

Activity #2 - Interval Recording

Interval recording of students' behavior has been found by many teachers to be a useful means of measuring the frequency of behaviors of concern, of measuring changes in students' behavior over time, and comparing the frequency of one student's behavior with that of his/her peers. Many teachers find that interval counting of students' behavior is a very useful task for teachers' aides to perform and that the recording of these observations serves as a helpful, on-going, individualized assessment strategy. As with other types of observation, experience with formal observation procedures also serves as good training for the informal observations that teachers are making every hour they are in the classroom.

For the purposes of direct observation, interval recording of behavior, a form of the Pupil Observation Schedule has been provided. Basic instructions for using that form are attached. Although relatively straightforward it should be impressed upon students that considerable familiarity with the category definitions, their placement on the form and the recording system itself will make the use of this instrument much more easily accomplished. It is also possible that some students might wish to develop their own observation instruments. This, of course, should be encouraged and the category definitions on the Pupil Observation Schedule should help them formulate definitions for their own categories.

PUPIL OBSERVATION SCHEDULE (Form B)*

Before beginning to observe, the observer should fill in name and date. Running notes about student characteristics, the setting and the activity can be taken during the observation, but the final statement should be written at the close of the observation.

All boxes (behavior descriptions) needed to describe the performance one or more times of a target behavior pattern by the student being observed during each 30 second observation period should be marked once with the appropriate symbol in the vertical column for that period. (The present form has 20 columns, enough for 10 minutes of observation of one student.) Use spaces between the lines on the form to write in descriptions of special behaviors noted if not otherwise coded.

Record all verbal and physical interactions as positive unless clearly negative in intent and/or effect on others.

Symbols: A check (✓) indicates the occurrences of a nondirected student behavior or, in the interactive categories, the occurrence of behavior involving either the teacher or the group as a whole including the teacher. For some categories, special symbols are used to record frequently occurring specific behavior patterns. These symbols are usually lower case letters. Use of these "special codes" is optional.

CATEGORY DEFINITIONS

ON TASK: Student is "attending to task" or "working." Eyes are directed toward task area. Task area can be away from desk and/or involve movement. "Work" could be a game or other activity. A student who is not attending to task may be "day dreaming," "playing," or engaging in some other activity. Try to make a narrative note on the specific non-task activity if time permits.

AT PLACE: Student is at teacher-approved place, usually at desk or table. Buttocks touching chair. Special codes: "T" when student turns head, shoulders or pelvis 90 degrees or more from "correct" task orientation for more than 4 seconds. "R" when student is rocking in chair so that one or more chair legs leave floor. A narrative note may be useful if "place" is away from usual work area. A student is away from a teacher-approved place when his or her buttocks are off the chair for more than 4 seconds. The behavior need not be "off task." Make a note if the student leaves the room and resume recording when a student returns.

POSITIVE VERBAL INTERACTION: Positive verbal interaction may be self-initiated or responding. Examples would be when students express verbal support for a peer or the teacher, ask constructive questions, give suggestions, offer ideas on topics being discussed, respond to teacher or peer questions, recite, and acknowledge help given by another. Special Codes: Mark "X" if a verbalization initiated is positive in intent but "comes at the wrong time." If time permits, note

*Original version (1973) by Frank H. Wood, Special Education Programs, Univeristy of Minnesota, Minneapolis, MN 55455. Revised 1979.

specific words and any positive gestures or expressions that accompany student's other behavior.

TEACHER INTERACTION: Teacher interaction with the target student is recorded as either positive or negative. This interaction may be verbal or nonverbal. Although the focus is on an individual student, records of the teacher's behavior directed toward the group of which the student is a part should also be recorded here. Task description and instruction, as well as verbal praise and encouragement, are to be recorded as positive (+). Criticism and threat are recorded as negative (-). Absence of marks in either of these categories during a single interval is an indication that there was no teacher interaction with the student during that time.

NOISE: Noise is recorded whether generated by the use of objects or the voice. Examples of object-generated noise are making a noise with the hands (drumming, pounding, clapping, etc.) or a hand-held object (pencil tapping, noisy scribbling on paper, crumpling paper, etc.) Vocal noise is when a student makes a noise with the mouth not directed specifically at others(humming, singing to self, shouting out, mumbling, whistling, etc.). Record the occurrence of noise if in the observer's judgement, the noise is audible to others in the group.

NEGATIVE VERBAL INTERACTION: Negative verbal interaction may be self-initiated or a response to statements by others. Examples of the first area would be student-initiated complaints about or criticism of the behavior of others. Also student verbal threats. Responding negative verbalizations are those made in response to statements or gestures by others or as part of a continuing dialogue or discussion.

NEGATIVE PHYSICAL CONTACT: Student hurts or interferes with the activity of another by touching him/her or his/her work or property. Student attacks another using hands, feet, or object either thrown or held in hand.

INDIVIDUAL BEHAVIOR: Space is provided where a record may be kept of an individual student behavior which is of special interest.

REMEMBER TO SUPPLEMENT SYMBOLS BY JOTTING IN A WRITTEN DESCRIPTION OF SPECIFIC BEHAVIOR OR THE WORDING OF A VERBALIZATION WHENEVER TIME PERMITS.

PUPIL OBSERVATION SCHEDULE (FORM B)

OBSERVER_____ DATE_____

DESCRIPTION OF STUDENT OBSERVED:

*DESCRIPTION OF OTHER PERSONS IN SETTING:

*DESCRIPTION OF SETTING AND ACTIVITY:

*(During this 10 minute period.)

		1	2	3	4	5	6	7	8	9	10
ON TASK		:	:	:	:	:	:	:	:	:	:
AT PLACE		:	:	:	:	:	:	:	:	:	:
POSITIVE VERBAL INTERACTION		:	:	:	:	:	:	:	:	:	:
TEACHER INTERACTION	(+)	:	:	:	:	:	:	:	:	:	:
	(-)	:	:	:	:	:	:	:	:	:	:
NOISE		:	:	:	:	:	:	:	:	:	:
NEGATIVE VERBAL INTERACTION		:	:	:	:	:	:	:	:	:	:
NEGATIVE PHYSICAL CONACT		:	:	:	:	:	:	:	:	:	:
_____		:	:	:	:	:	:	:	:	:	:

Suggested Readings

Boehm, A. E., & Weinberg, R. A. The classroom observer: A guide for developing observation skills. New York: Teachers College Press, 1977.

Deno, S. L., & Mirkin, P. K. Data-based program modification: A manual. Reston, VA: Council for Exceptional Children, 1977.

Walker, H. M. The acting-out child: Coping with classroom disruption. Boston: Allyn and Bacon, 1979.

Weinberg, R. A., & Wood, F. H. (Eds.). Observation of pupils and teachers in mainstream and special education settings: Alternative strategies. Reston, VA: Council for Exceptional Children, 1975.

In R.B. Rutherford & A.G. Prieto (Eds.), Severe behavior disorders of children and youth. Monograph in behavior disorders, Summer, 1979. Reston, VA: Council for Children with Behavior Disorders, Council for Exceptional Children, 1979.

NORMATIVE BEHAVIORAL OBSERVATION DATA AS A STANDARD IN CLASSROOM TREATMENT OF EDUCATIONALLY HANDICAPPED CHILDREN

Steven R. Forness

ABSTRACT

Although marked discrepancy between a child's behavior and that of his or her classmates is often a reason for initiating treatment, little empirical data is available upon which to base such a decision. Observations of 220 children in 19 classes for the educationally handicapped were made over several days to provide a tentative normative standard for such classrooms. Means and standard deviations were examined in four categories, along with teacher and peer responses to each type of behavior. Relatively low frequencies of disruption were found along with relatively high levels of teacher attention to on-task behaviors.

Discrepancy between a child's behavior and that of his or her classmates has frequently been cited as a reason why teachers either refer the child for treatment or initiate interventions designed to change the child's behavior in the classroom setting (Bolstad and Johnson, 1977; Forness and Esveldt, 1974; Nelson, 1971; Patterson, Cobb, and Ray, 1972). Although direct observation is then used to establish a baseline frequency of problem behavior, and to evaluate effects of treatment (Forness, 1970; Strain, Cooke, and Apolloni, 1976), systematic consideration is rarely given to a child's behavior relative to other children in the classroom. Only recently have investigators begun to employ a system of sampling the behavior of peers as a method of evaluating classroom intervention procedures (Patterson, 1974; Walker, Hops, and Johnson, 1975; Walker and Hops, 1976).

It has been suggested that normative data on peers is critical both in monitoring change in a particular child's behavior, as well as in deciding the "normal" limits of behavior for a given classroom setting (Forness, 1975); Greenwood, Walker, and Hops, in press; Walker and Hops; 1976). While a few investigators have observed behavior of children in regular classrooms in this regard (Bryan, 1974; Forness and Esveldt, 1975 a and b; Gottman, 1977; Nelson, 1971; Richey and McKinney, 1978; Werry and Quay, 1969), relatively little data is available on observable behavior of children
94 in special class settings, despite the fact that children in these settings

continue to require treatment for behavior problems (Barr and McDowell, 1972; Fink, 1972; Haubrich and Shores, 1976). As has been suggested in other areas of assessment (Gunzburg, 1973; Mercer and Lewis, 1977), it is often helpful to have a frame of reference which also includes normative data for "special" populations.

The present study describes expected levels of observable classroom behavior for children in a relatively representative sample of classes for the educationally handicapped. These classes contain both learning disabled and emotionally disturbed children and represent the special class placement most commonly used in California for children with behavioral problems (Hansen, 1970; Keogh, Tchir, and Windeguth-Behn; 1974). The observation technique used was developed from previous work on classroom intervention (*c.f.*, Forness, 1975), and includes behavioral categories relevant to intervention procedures, such as off-task behavior and classroom disruption. The technique has also been used to describe behavior of educationally handicapped children in regular elementary classroom settings (Forness and Esveldt, 1974, 1975a), and to identify young children at risk for subsequent educational handicaps (Forness and Esveldt, 1975b; Forness, Guthrie, and Nihira, 1975; Forness, Guthrie, and Hall, 1976; Forness, Hall, and Guthrie, 1977).

METHOD

Participants and Settings

There were 220 children observed in 19 classrooms for the educationally handicapped selected from among three different counties in Southern California. These children represented 1.2 percent of all children the same age enrolled in such classes in California. School districts represented as closely as possible the demographic mix in large and small districts, urban and suburban communities. Specific breakdown, however, was unavailable on ethnic minority variables.

Class size ranged from nine to 14 with a mean of 11.5 children per classroom (S.D. = 1.9). Age of the children ranged from eight to 13 with a mean of 10.7 (S.D. = 1.4). Seventy percent of the samples were males. All participants were observed while enrolled in educationally handicapped classes during the spring of 1977. As indicated above, all had been referred to these classrooms because of learning disabilities and/or behavior problems in the regular class.

Observation System

Behaviors were recorded in four categories: *Verbal interaction*—defined as task-oriented verbal or gestural attempts to communicate, such as asking questions, reciting, or raising hand; *on-task behavior*—defined as eye contact to teacher, task materials, or peer who is reciting; *off-task behavior*—defined as eye contact to other than above; and *disruptiveness*—defined as behavior incompatible with on task activities, such as talking to peers when not permitted, speaking out of turn, throwing objects, verbal or physical aggression, etc. Categories were treated as mutually exclusive, *i.e.*, only one could apply during an interval.

Behaviors were also recorded under one of three conditions. If the teacher or a peer happened to be attending or responding directly to the child during the interval, the behavior was recorded under *teacher response* or *peer response*; otherwise, behavior was recorded under *no response*.

A time-sampling procedure was used involving six-second intervals marked off with red tape on a stop watch attached to the observer's clipboard. All children in each classroom were observed in round-robin fashion according to their order of appearance on the data sheet. During each six-second interval, the observer located the next child on the sheet, observed long enough to form a mental image of the behavior, placed a tally by the appropriate behavior under the corresponding response condition, and went on to the next child on the sheet. Observations continued until 10 rounds had been made on the classroom for the daily observation period.

Six different observers were trained in two group sessions lasting a total of six hours. Both paper and pencil reliability exercises and simulation conditions were used. Particular emphasis was placed on problems of reliability and observer bias (Mash and McElwee, 1974; Taplin and Reid, 1973; Wahler and Leske, 1973). Observer reliability was checked by two observation supervisors previously trained to criterion over several sessions by the author. Reliability checks were made during the first two weeks of data collection. Supervisors recorded data simultaneously with each observer on the same group of children. Reliability coefficients were computed by dividing agreements by the total of agreements plus disagreements. These averaged .89 with a range of .83 to .98 for all observers.

Procedures

All classrooms were observed, as much as possible, during the morning hours at a time when all children were either functioning as a group or were at least engaged in a similar type of activity, such as seatwork. Before observations were begun, each observer spent at least one day in the classroom learning children's names from a seating chart, becoming accustomed to classroom routines, and doing a brief "practice round" of observations which was not counted in the data. This allowed time for the children and teacher to get used to the observer's presence, as suggested by Masling and Stern (1964).

An observer continued observing an assigned classroom until that particular classroom had been observed a minimum of four days, a period suggested as necessary for a stable measure of children's classroom behavior (Forness and Guthrie, 1977). Although all children were observed for the same amount of time each day, they were observed for varying number of days so their individual totals in each category were converted to percentages.

RESULTS

Mean percentages of behavior in each category and response condition are presented in Table 1, along with standard deviations. Subjects engaged in appropriate classroom behavior, the total of verbal interactions
96 plus on-task behavior, 82.8 percent of the time (S.D. = 14.3 percent), rang-

TABLE 1
Mean Percent of Behavior Observed Under Three Response Conditions

	No Response Mean (S.D.)	Teacher Response Mean (S.D.)	Peer Response Mean (S.D.)	Total Mean (S.D.)
Verbal interaction	1.6 (3.4)	9.9 (10.6)	1.9 (4.2)	13.4 (12.0)
On-task behavior	63.4 (18.3)	3.6 (4.5)	2.4 (5.0)	69.4 (16.2)
Off-task behavior	11.5 (10.8)	0.4 (1.5)	1.5 (3.1)	13.4 (11.9)
Disruptiveness	1.8 (3.8)	0.8 (2.1)	1.1 (2.8)	3.7 (6.4)
Total	78.3 (13.8)	14.7 (11.5)	6.9 (7.3)	100.0

ing from 21 to 100 percent. The level of disruptiveness ranged from 0 to 47 percent. The mean time which teachers were observed in response to children's total appropriate behavior was 13.4 percent (S.D. = 11.4), ranging from 0 to 60 percent. For peers, this same figure was 4.3 percent (S.D. = 6.8), ranging from 0 to 50 percent. For total inappropriate behaviors, the response ranged from 0 to 23 percent for teachers and from 0 to 24 percent for peers.

The only significant sex difference was that boys were more disruptive than girls under the "no-response" condition (unpaired t test with unequal variance, $t = -2.83$, $p < .01$); but the actual percent difference was negligible. There were not apoarent relationships between total appropriate behavior and class size nor between this behavior and age ($rho = .04$ and $-.14$, respectively).

DISCUSSION

The findings seem to suggest that any child whose percentage of appropriate behavior (verbal interaction plus on-task) falls much below 70 percent begins to be at risk for behavior problems. It is when that level begins to reach 50 percent that intervention would seem warranted. Depending on the case, figures on disruptive behavior may be more to the point. A child whose disruptiveness is over 10 percent becomes at risk, and any child much over 15 percent becomes a candidate for intervention. Some caution should be used here, however, since the disruptive category provides no distinction between behaviors of high intensity but low frequency, e.g., aggressive or assaultive behavior, and disruptions which are seen as less immediately troublesome (Geston, Cowen, DeStefano, and Gallagher, 1978). These figures are of course, limited to educationally handicapped classes as defined herein.

How these figures are used from classroom to classroom with individual educationally handicapped children is another matter. There were some classrooms in the sample where literally no children were even in the "at risk" range and one classroom where *two* of the 10 children appear to

need immediate intervention. The question of treatment is always a relative one; but normative data in this case might assist a teacher, for example, in pressing his or her case for additional help in the classroom or referral for treatment. Effectiveness of subsequent treatment or intervention could then be given additional credence, beyond the child's improvement over his or her own baseline, whenever levels of appropriate behavior begin to rise substantially above 70 percent.

Another monitoring point could be levels of teacher or peer response. For example, whenever a child appears to receive attention more than 10 percent of the time from peers for inappropriate behavior, contingent use of peer attention and ignoring (*c.f.*, Patterson, 1974) might be seen as the treatment of choice. Unfortunately, these data do not seem to be particularly useful for identifying socially isolated children (Gottman, 1977; Greenwood, Walker, and Hops, in press) since no peer response at all still appears to be in the normal range. It is interesting to note the apparently effective distribution of teacher attention in which teachers appear to respond 10 times more often to appropriate behavior than they do to inappropriate behavior.

Another interesting finding is that previous research with the same observation technique (Forness and Esveldt, 1974a) indicates that young educationally handicapped children undergoing referral for special classes, while still in the *regular* classroom, were more than 10 percent below children in the present sample in levels of appropriate behavior. An unexpected finding is that levels of children's appropriate behavior in educationally handicapped classrooms tend to equal that found for normal children in regular classrooms (Forness and Esveldt, 1974; Forness, Guthrie, and Nihira, 1975). Whether these two findings can be considered a testimonial to special class placement is purely conjecture. The use of normative observation data does appear, however, to bring an additional perspective to behavioral treatment of children in special classroom settings.

REFERENCES

Barr, K. L. and McDowell, R. L. Comparison of learning disabled and emotionally disturbed children on three divant classroom behaviors. *Exceptional Children*. 1972, *39*, pp. 60–62.

Bolstad, O. D. and Johnson, S. M. The relationship between teacher's assessment of students and the student's actual behavior in the classroom. *Child Development*, 1977, *48*, pp. 570–578.

Bryan, T. An observational analysis of classroom behaviors of children with learning disabilities. *Journal of Learning Disabilities*, 1974, *79*, pp. 191-203.

Fink, A. H. Teacher pupil interaction in classes for the emotionally handicapped. *Exceptional Children*, 1972, *38*, pp. 469–474.

Forness, S. R. Behavioristic approach to classroom management and motivation. *Psychology in the Schools*. 1970, *7*, pp. 356–363.

Forness, S. R. Looking: Use of classroom observation in early identification. In Rutherford, R. and Buckhalt, J. (Eds.) *Distinguished lectures in special education*. Los Angeles: University of Southern California Press, 1975, pp. 19–41.

Forness, S. R. and Esveldt, K. Classroom observation of children referred to a child psychiatric clinic. *Journal of Child Psychiatry*. 1974, *13*, pp. 335–343.

Forness, S. R. and Esveldt, K. Classroom observation of learning and behavior problem children. *Journal of Learning Disabilities*, 1975, *8*, pp. 382–385 (a).

Forness, S. R. and Esveldt, K. Prediction of high-risk kindergarten children through classroom observation. *Journal of Special Education*, 1975, *9*, pp. 375–387 (b).

Forness, S. R. and Guthrie, D. Stability of short-term classroom observations. *Psychology in the Schools*, 1977, *14*, pp. 116–120.

Forness, S. R., Guthrie, D. and Hall, R. Follow-up of high-risk children identified in kindergarten through direct classroom observation. *Psychology in the Schools*, 1976, *13*, pp. 45–49.

Forness, S. R., Guthrie, D. and Nihira, D. Clusters of observable behavior in high-risk kindergarten children. *Psychology in the Schools*, 1975, *12*, pp. 263–269.

Forness, S. R., Hall, R. and Guthrie, D. Eventual school placement of kindergartners observed as high risk in the classroom. *Psychology in the Schools*, 1977, *14*, pp. 315–317.

Gesten, E., Cowen, E., DeStefano, M. and Gallagher, R. Teachers' judgements of classroom-related teaching-related problem situations. *Journal of Special Education*. 1978, *12*, pp. 171–181.

Greenwood, C. R., Walker, H. M. and Hops, H. Issues in social interaction/withdrawal assessment. *Exceptional Children*, (in press).

Gottman, J. M. Toward a definition of social isolation in children. *Child Development*, 1977, *48*, pp. 513–517.

Gunzburg, H. C. *Social competence and mental handicap*. Baltimore: Williams and Wilkins, 1973.

Haubrich, P. A. and Shores, R. E. Attending behavior and academic performance of emotionally disturbed children. *Exceptional Children*, 1976, *42*, pp. 337–338.

Hansen, P. A. Children called E. H. *Journal of California Council for Exceptional Children*, 1970, *20*, pp. 5–12.

Keogh, B., Tchir, C. and Windeguth-Behn, A. Teacher's perceptions of educationally high risk children. *Journal of Learning Disabilities*. 1974, *7*, pp. 367–374.

Mercer, J. R. and Lewis, J. F. *System of Multicultural Pluralistic Assessment*. New York: Psychological Corporation, 1977.

Masling, J. and Stern, G. Effect of the observer in the classroom. *Journal of Educational Psychology*, 1964, *60*, pp. 351–354.

Mash, E. J. and McElwee, J. D. Situational effects on observer accuracy: Behavioral predictability, prior experience, and complexity of coding categories. *Child Development*, 1974, *45*, pp. 367–377.

Nelson, C. M. Techniques for screening conduct disturbed children. *Exceptional Children*, 1971, *37*, pp. 501–507.

Patterson, G. R. Intervention for boys with conduct problems: multiple settings, treatments and criteria. *Journal of Consulting and Clinical Psychology*, 1974, *42*, pp. 471–481.

Patterson, G. R., Cobb, J. A., and Ray, R. S. Direct intervention in the classroom: A set of procedures for the aggressive child. In F. W. Clark,

D. R., Evans, and L. A. Hammerlynck (Eds.) *Implementing behavioral programs in educational and clinical settings*. Champaign, Illinois: Research Press, 1972.

Richey, D. and McKinney, J. Classroom behavioral styles of learning disabled boys. *Journal of Learning Disabilities*, 1978, *11*, pp. 297–302.

Strain, P., Cooke, T., and Apolloni, T. *Teaching exceptional children: Assessing and modifying social behavior*. New York: Academic Press, 1976.

Taplin, P. S. and Reid, J. B. Effects of instructional setting and experimenter influence on observer reliability. *Child Development*, 1973, *44*, pp. 547–554.

Wahler, R. G. and Leske, G. Accurate and inaccurate observer summary reports. *Journal of Nervous and Mental Disease*, 1973, *156*, pp. 386–394.

Walker, H. M. and Hops, H. Use of normative peer data as a standard for evaluating classroom treatment effects. *Journal of Applied Behavior Analysis*, 1976, *9*, pp. 159–168.

Walker, H. M., Hops, H., and Johnson, S. M. Generalization and maintenance of classroom treatment effects. *Behavior Therapy*, 1975, *6*, pp. 188–200.

Werry, J. S. and Quay, H. C. Observing the classroom behavior of elementary classroom children. *Exceptional Children*, 1969, *35*, pp. 461–470.

Steven R. Forness, Professor, UCLA Neuropsychiatric Institute, Los Angeles, California.

In R.A. Weinberg & F.H. Wood (Eds.), Observation of pupils and teachers in mainstream and special education settings. Reston, VA: Council for Exceptional Children, 1975.

An Historical Overview of Systematic Approaches to Observation in School Settings

Richard M. Brandt
University of Virginia

In striking contrast to a 1960 report in which only 8% of the empirical studies in child development literature appeared to be based on observational data (Wright, 1960), recent educational and behavioral science writings contain many observational researches. Ethologists are now applying their animal-oriented techniques to the study of human behavior. Periodicals like the *Journal of Applied Behavior Analysis* have been introduced to provide publication outlets for behavior modification studies and single subject research. After years of laborious work making specimen records of child functioning in natural environments, ecologists are attracting an increasingly broad audience of educators interested in both their findings and procedures. Would-be investigators now can select interaction scales from a vast array of possibilities. The 73 classroom scales reproduced only recently in the *Mirrors for Behavior* collection (Simon & Boyer, 1967, 1970) no longer represent even a majority of the instruments available for studying teacher behavior, not to mention numerous scales for child behavior and other observation targets.

The current and future trends stressed in this conference stem from deep, historical roots. The writings of Darwin, G. Stanley Hall, Gesell, Lorenz, Tinbergen, Dorothy Thomas, Harold Anderson, Robert Bales, Herbert Thelen, and Hartshorne and May spring immediately to mind, as do the early anthropological studies of primitive cultures by the likes of Boaz, Benedict, and Mead, and of community life by Warner, Hollingshead, William Foote Whyte, and many others. In education, the work of Daniel Prescott and his staff at the Universities of Chicago and Maryland perhaps highlight the child study movement in America. The contributions have been continuous and additive over two and three decades for people like Roger Barker, Donald Medley, Jacob Kounin, Ned Flanders, and Phil Schoggen; and over five decades for Jean Piaget, who published his first works based on child observation in the early 1920's and whose child development

BEST COPY AVAILABLE

perspective continues to have profound impact on educational theory and practice.

Nor, in an historical overview, should we neglect to mention parallel traditions in the use of observational procedures by industrial and commercial institutions to study time and motion of manufacturing operations and participatory management practices. Here the names of Roethlisberger, Gilbreth, Bales, and William Foote Whyte come to mind as among the early pioneers.

Although the early work of each of these individuals, as well as that of many others, is worthy of careful study for its methodology and lasting contributions, in the interest of proper keynoting let me concentrate my remarks on the general nature of the various observational approaches which will be presented at this conference and the similarities and differences among them. I hope to provide an overall framework for seeing the relations and potential contributions of each of these approaches to the study of educational activity, but my main purpose is to open up issues and make beginning and tentative generalizations about particular substantive and methodological questions as a basis for further exploration in the following sessions. Given the scope of the conference and the caliber of talent represented in the other presentations, I cannot expect to do more than deal superficially with the many matters that deserve consideration. The conference schedule allows for increasing depth in the presentation of each perspective, and I am counting on other presenters to set me straight if I misconstrue their positions or gloss over what they consider important matters.

Let me then, without worrying about giving adequate and full coverage to each of the positions or all of the potential issues, address myself first to some of the reasons why educators are turning increasingly to observational methodology and the kinds of questions that they would hope to answer via its employment.

Reasons for Classroom Observation Studies

Education is often thought to be more art than science. Many educators are skeptical of attempts to analyze instruction objectively. They consider the search for an empirical base for our profession merely another educational pipe dream. I stand in firm opposition to this skepticism; and, without minimizing the difficulties to be faced in pursuing such an objective, I feel that not only is the task surmountable but clearcut progress is already apparent.

Dimensions of Educational Processes

Perhaps the most pervasive and global question that educators persistently ask is what are the critical dimensions of school life in shaping child learning and development. Which of the many variables we might measure are most likely to account for whatever behavior changes we seek? The outcomes we choose to concentrate on represent value decisions, but we need to know which process variables best correlate with such outcomes once we make these decisions.

Obviously, some analysts believe that the school makes only a limited total contribution to overall developmental patterns. The writing of Jencks, Coleman, and Bloom, among others, do indeed suggest that other socializers than the school—socioeconomic status, home relationships, entering characteristics, and peer factors particularly—are highly important by comparison. Stearns' (1971) review of Head Start evaluative studies indicates that the ability and performance gaps between advantaged and disadvantaged children are seldom diminished by attendance in such programs. Medley (1973) recently concluded, after reviewing numerous studies of teacher effectiveness in relation to pupil achievement gains, that almost no single teacher characteristic has been discovered that consistently makes a significant difference in what a child learns.

Despite these disappointing beginnings, the search to identify those dimensions that do make a difference goes on; and there are, it seems to me, some promising leads. In several expertly staffed experimental programs designed around clearly specified educational objectives for guiding daily activities and utilizing sufficient instructional resources and parent involvement, significant short-term learnings have been found in various cognitive and perceptual skills (Stearns, 1971, p. 161). Using observational data for assessing program differences between various Follow Through models, Stallings (1974b) recently reported that although entering ability accounted for slightly over half of the total variance in child outcome measures, instructional processes alone accounted for 26% of the variance for one set of data from 30 first-grade classrooms and 9% of the variance in another sample of 112 first-grade classrooms. In 58 third-grade classrooms, furthermore, instructional process variables alone accounted for more of the variance than entering ability in mathematics achievement scores and for about the same amount of variance on the Ravens Progressive Matrices Test. The overall amount of child outcome variance accounted for by the entering ability, instructional process, and interaction variables was 71% and 62%, respectively, for the two samples. The works of both Soar (1972b) and Flanders (1970) like-

wise have shown consistent significant relations between teacher behavior and instructional style, on the one hand, and student learning and change, on the other.

Student Outcomes

Before saying more about what specific dimensions seem to be important, let me mention a second major usage for observational methodology besides the identification and measurement of key instructional and classroom processes. I am referring to the assessment of student outcomes on many cognitive and affective variables that traditional tests do not measure very effectively. I am sure that one of the main reasons that educators are turning to observational methodology more and more is their increasing disenchantment with standardized tests as the primary vehicles for dependent variable assessment. Even if some of these test procedures deserve good marks on the basis of high reliability and validity coefficients for assessing particular abilities, they fall far short of providing us with the tools needed to measure most of our educational goals or to permit comprehensive assessment of program effectiveness. I do not mean to suggest that we stop using standardized tests but that we should supplement their usage with solid measures of those aspects of pupil learning and behavior for which adequate tests are not available. Test-taking is but a small fraction of the responses children make in or out of school; and we need to be recording accurately, with the help of systematic observation, what some of their other responses are to ordinary life stimuli.

If achievement tests leave something to be desired in ability assessment, the self-report scales and projective tests typically used in assessing self-concept, locus of control, anxiety, and other affective dimensions are even less exacting. Relations between test responses and actual feelings or status are often obscure. The subtle, defensive workings of human personality usually preclude fully valid assessment by means of such tools.

Systematic observation of behavior under natural conditions, however, provides an acceptable alternative, or at least supplement, for such assessment, if data are sufficiently complete and objective to permit various response patterns to be identified precisely in relation to contextual changes. The list of child behavior patterns it is possible to identify via observational means is almost limitless. Illustrative of child variables already apparent in child development literature are such items as "child expressing positive affection and approval to others," "child approaching or waiting for others," "attending to

task," "child-initiated interaction," "asking open-ended questions," "making higher level cognitive responses," "sorting materials on a relational category basis," "being distracted," "biting fingernails or having something in mouth," "adhering or not adhering to specific school rules," "being absent or not absent," and many other categorical type behaviors of interest. Two of my students have been working most of this semester to develop an interactional, event-sampling checklist that permits an observer to keep a running record of preschool children's follower-leader behavior during free play. Such a checklist would permit us to record who follows whom and how the leadership is manifested in particular verbal and nonverbal categories. Using this checklist, we should be able to determine the stability of leadership-followership behavior over time and across settings.

Process-Outcome Relations

The potential utility of observational process and outcome measures lies primarily in the linkage between them. The ultimate criteria for identifying educational dimensions of the greatest importance should be those that not only distinguish one classroom or program from another but also account for significant variance in child outcome measures. As indicated earlier, we are beginning to find consistent relations of this sort showing up in some of the Follow Through research reports. Both Soar (1972a) and Stallings (1974a) have reported significant and sometimes rather high relations between process and outcome variables, though not always in clearly predictable fashion from one program model to another, nor are the relations always linear. Soar found that pupil gain on achievement tests was typically greater in classrooms with a moderate amount, rather than too little or too much, of pupil freedom and teacher control; and the relations varied, furthermore, with the type of learning under investigation.

Turning to behavioral measures of child outcome, I was impressed at a recent AERA convention by the findings of several Follow Through studies. A series of papers was presented showing that children who attend open education programs, when compared with those in either other Follow Through programs or traditional classrooms, exhibit greater independency, higher-level cognition, more cooperation, and lower absence rates (Goldupp, 1974; Ross & Zimiles, 1974; Stallings, 1974b). One of the presenters reported that by the middle of the academic year, children exhibited less inappropriate behavior when the teacher was out of the room than had been displayed in the fall; third graders, in the fall, displayed less

inappropriate behavior than first graders did in the spring. Children in "structured programs," furthermore, displayed more task persistence. In brief, solid evidence was forthcoming from observational data alone that some of those hard-to-measure, affective objectives of many educational programs were being met. Observation offers the promise, therefore, of not only assessing both process and outcome separately but also discerning relations between the two and ultimately helping to determine what are and what are not significant educational dimensions.

Overview of Four Perspectives

Ecological Psychology

Let me turn now to the several observational perspectives (for want of a better word) and some discussion of the kinds of educational problems for which they might be best suited. First the ecologists.

We have been reading the books and articles of the Barker group, including Phil Schoggen and Jacob Kounin, for a quarter of a century. We admire their dogged pursuit of the parameters of real-life phenomena and their seemingly hopeless attempts to describe accurately and completely the totality of behavior and the surrounding complex of interacting forces that help shape it. Rooted in early Gc talt psychology, the ecologists are not satisfied to accept the narrow dictates of laboratory research with its tight controls and precise measuring sticks but' set about, instead, to capture with as much scientific rigor as possible the essence of ordinary behavior under natural conditions. One has to be impressed with their painstaking attempts to describe fully and accurately the sequence of events in one boy's day and the activities and happenings in small towns in the Midwest, Yorkshire, England, and other areas. The movement as a whole has provided the most extensive set of narrative records of ordinary life behavior in existence.

For many years this group seemed to be like voices in the wilderness. Psychologists were aware of their work, but their notion was not generally accepted that one could truly reach an understanding of the operations of psychological forces without careful manipulation and control of psychological variables under relatively controlled conditions. Ecologists presumably dealt with too many variables at one time, too imprecise a means for assessing them, and too limited samples of people and events for adequate generalizations to be made. Atheoretical in their overall approach, ecologists were seldom able to confirm or deny their theoretical propositions and constructs. They

certainly could not be subjected to rigorous testing within the confines of their research methodology.

Currently, however, the painstaking labors over the years are being rewarded and the voices listened to by an increasingly wider audience. The Willems and Raush book, *Naturalistic Viewpoints in Psychological Research* (1969), and Barker's *Ecological Psychology* (1968) have both attracted considerable attention. Herbert Wright's elaborated description of the narrative recording methodology used in *Midwest and its Children* (Barker & Wright, 1955) has been published in paperback (1967), an indication of popularity. In education, the findings reported in *Big School, Small School* (Barker & Gump, 1964) often have been cited by school administrators who are engaged in the process of deciding the kind of new school a town needs. The mere inclusion of representatives from this perspective in a conference such as this one is indicative of its recognized contributions to observational methodology.

The fundamental assumption underlying ecological psychology probably is the notion of an intrinsic order in human events and behavior that will occur with consistency and regularity within and in response to the surrounding environmental forces, whether or not our theories happen to take notice of this order (Gutmann, 1969). The task of the observer is to gather sufficient data to uncover some of this order and behavioral consistency, and to discern particularly how behavioral patterns change as they interact with setting and contextual forces.

Sechrest (1969), among others, has noted the inconsistencies between human behavior under testing and laboratory conditions, on the one hand, and human actions in ordinary life situations, on the other. In partial response to this inconsistency, social psychologists in their recent laboratory researches have been attempting to simulate more closely than in earlier research relatively complex and realistic life-like conditions.

The ecologists, furthermore, have continued to study natural behavior in ordinary settings, but with increasing rigor and more solid comparative data across studies from one setting to another. They have even used time-lapse photographic methods to objectify some of their data. The slow acceptance of the contributions from the ecological movement is due, in part, to the fact that sufficient data have been available only recently for useful comparisons of behavioral data from one place and group to another. For example, Schoggen, Barker, and Barker (1963) were able to report that the behavioral units of Midwest children were of shorter duration on the average than those of comparable English children, thus lending support to

the often-cited notion that English children are less hyperactive than American children.

In addition to providing detailed specimen records of the behavior of individuals, ecological psychologists have analyzed the environmental forces existing within and among communities and institutions. Thus, the ecologists continue to pioneer the development of strategies for understanding individual behavior in its environmental context.

Another special contribution of ecological psychology to our understanding of human behavior is in generating hypotheses worthy of more carefully controlled experimental research. Currently, psychology has barely begun to explain daily happenings and the ordinary behavior of people within the complex network of ongoing circumstances. We barely know how to focus on the vast array of interacting variables as we try to account for what people do and say.

I referred earlier to the need for identifying in school life the key dimensions from the many that exist. It seems to me that specimen records and anthropological reports of daily school life, such as those kept by Louis Smith all day long over a whole semester (Smith & Geoffrey, 1968), are badly needed to help identify promising dimensions that have not been considered previously. The profound impact that Phillip Jackson's book, *Life in Classrooms* (1968), has had on many educators is due in part to the identification of certain dimensions of school settings that had not previously been recognized, at least not with such convincing observational data on the prevalency of their occurence.

A basic feature of ecological studies is that they are primarily atheoretical. The observer enters the study situation with relatively few preconceptions of what he will find and, as a consequence, he is more alert to discover previously unrecognized happenings. He records as much and as fully as he can all that is happening. His aim is to freeze complex behavioral events so he can examine the "stream of behavior" more reflectively and systematically later. His analysis is inductive and heuristic; first he breaks up events into codable units—what Barker referred to as dividing the behavior stream—and then he looks for patterns of behavior in specific settings and under particular circumstances. As he generates hypotheses for particular patterns, he is able to code his records deductively for further supporting or negating evidence. His records are basically descriptive, but the inferences he derives become the basis not only for sharpening educational awareness but also for solid theory building and eventual experimental testing.

The method is particularly appropriate for the times we want to explore certain aspects of school life to see what they look like under

close scrutiny and we are relatively open minded about what we might find. Perhaps the most insightful and useful bit of data I obtained during a three-week study of a British infant school came from an audio tape recording of a one-hour classroom show-and-tell discussion. After three weeks of data gathering with one of Donald Medley's interaction instruments, I had become quite fascinated by the manner in which teachers questioned children and caused them to reflect on their activities in an open classroom; but my casual observations of how this questioning was done and the precise form it took had truly escaped me. It was not until I had the chance to listen several times to this tape of a typical show-and-tell discussion that I was able to define and then count the specific types of questions and comments the teacher made, in order to describe her discussion leadership style explicitly. Although it is not a specimen record, this tape represented the type of "frozen" behavior the ecologists work with, and the analysis procedures were quite similar. The results, furthermore, were highly rewarding because the particular types of teacher questions and comments that were being used did not fit any category sets or theoretical descriptions that I had previously seen (Brandt, 1973).

If I paint the work of the ecological psychologists in very flattering colors, I do so because I firmly believe that they have much to offer the educator. However, I should not leave the impression that they have all the answers and the other perspectives can be dismissed. Since their procedural problems are numerous, at least a few should be mentioned. Their work is tedious and slow. The dross rate of unused information that is recorded and leads nowhere by way of substantial findings is high. Relatively few clearcut implications of high educational value can be drawn from their findings at this time, considering the years of work the ecologists have put in. The amount of work that must be done before generalizations can be achieved with any certainty is great; most findings must be put to further empirical testing through replication or experimental research before generalizations would seem warranted.

One particular problem centers on the question of how to divide the behavior stream; another, on what molarity level to code. It is often very difficult to tell when one behavior stops and another starts or even what is a distinct behavior unit to code and what is only a subunit. The distinction between molar and molecular behavior is not always clear, nor are specimen record data always as objective as they might be. Including inferential language as part of the recorded data to be analyzed presents special problems for reliability assessment and interpretation. Having spent over a dozen years training

teachers to write and analyze good naturalistic case records of children for inservice learning purposes, I have great faith in the soundness of the general methodology but recognize well its limitations. I hope that the representatives of this approach will be able to address some of the concerns I have raised and show us more clearly than I have what its contributions to educational thinking are or might be in the future.

Ethological Research

Closely related to the ecological approach is that of the ethological approach. I shall make fewer remarks about what we can look forward to from this perspective because I know less about its followers. As I understand their work, those who are concentrating on human research are attempting to use essentially the same methodology that has been used in studying animal behavior in its natural habitat. Direct, precise, and complete observation is its essential methodology. As Carthy stated (1966), "The first aim of the study of the behavior of a particular animal is to record it in all its details, correlating it with the stimuli which evoke the different sections of it. Such a complete catalogue of behavior is called an ethogram" (p.1).

Hutt and Hutt (1970, pp. 22-23) identified four characteristics that distinguish the ethological from the ecological approach.

1. Whereas the ecologist concentrates attention primarily on molar actions involving the total person, such as going to school or playing baseball, ethologists focus on smaller action units such as facial expressions, gestures, or visual fixations. The actones or molecular behaviors of the ecologists, performed by muscle groups, represent the *data desiderata* for the ethologists, who seek a more microscopic analysis of behavioral change and stability amid shifting environmental contexts.

2. In dividing behavior protocols into episodes, the ecologist infers the goals of the actor, which requires continual analysis at different stages in the transcript. The ethologist avoids this rather arduous, inferential, and somewhat uncertain process by ignoring the actor's goals and delineating the units of behavior by his motor patterns alone.

3. Ecologists typically include inferences of attitudes, motives, and intentions in their specimen records; ethologists specifically exclude human thoughts and desires from their reporting. Whereas the ecologist attempts to describe the "psychological habitat" or life space, to use the earlier Gestalt phrase for the environment as viewed by the actor, the ethologist considers such reporting unscientific.

4. In stressing how a subject does something, the ecologist uses numerous adjectival and adverbial terms and phrases to characterize the quality of the actions. Specimen records contain such expressions as "Looking wise and kind, Ben..." and "Roy stared after Bill with a hurt and hostile look." Ethologists would not accept such ambiguous, nonbehavioral descriptions in their records.

In contributing to a science of human behavior, therefore, the ethologist accepts as his first task the accurate, objective description of man's behavior under natural conditions. His emphasis on objectivity leads him to focus on motor patterns that can be observed clearly and discretely so that one cannot doubt the occurrence of the events he reports nor the precise manner of their occurrence. As Hutt and Hutt (1970) stated,

> It is only by repeated sampling of a child's behavior in many different situations that the consistencies in behavior emerge. Whereas on initial contact the child's behavior appears to be infinitely variable, with repeated observation it becomes clear that certain *patterns* tend to recur in similar circumstances, that patterns bear a temporal relationship to each other, and that some patterns occur frequently, others infrequently. Most important of all, we begin to realize that far from being infinitely variable, the child's repertoire of behaviors is finite. (p. 29)

Although the history of ethology goes back at least to Charles Darwin and there have been regular contributions by ethologists throughout the first half of this century, especially for nonhuman species, the specific study of man's behavior by ethological methods has only recently gathered momentum. The popularized works of Lorenz and Tinbergen have been a stimulating force behind some of the current interest, no doubt. However, the writings and reports of some of the current researchers are more important. Cross-species and cross-cultural comparisons are showing up in the behavioral science literature with increasing frequency, and the relevance of ethological ideas and methods to our understanding of child behavior in particular is becoming ever more apparent.

An important recent contribution is McGrew's (1970) glossary of 133 motor patterns exhibited by young children in an English nursery school. Among the 20 most prevalent patterns are *turn* (i.e., to rotate the trunk face first about the body's longitudinal axis), *run* (i.e., to move the body rapidly forward, alternating legs during each stride, so that both feet are simultaneously off the ground during each stride), and *reach* (i.e., to move the arm toward an object by arm extension and pronation, finishing palm down with hand open and fingers spread). The glossary gives the number of observations and mean

duration of each pattern. Also cited are general movements of motor patterns: hand, arm, leg, gross body, other or locomotory.

Ethological studies report the following kinds of findings:

1. Inexperienced nursery-school children, compared to experienced children, jumped and skipped less, walked slowly more often, vocalized loudly less frequently, verbalized less, switched from one toy to another more frequently, spent more time monitoring adults' activities, initiated interaction with adults more often, spent more time in close proximity to them, and allowed themselves to be directed by experienced children more than vice versa. In contrast to near-human primate group formation patterns, aggressive behavior of young children was initially infrequent and gradually increased over the beginning seven days of nursery school (McGrew, 1972, pp. 134-148).

2. When compared with their behavior when their mothers were present, the speech, movement, and play of infants aged 13-15 months decreased after their mothers left the room. The frequency of each of these behaviors showed a slight but progressive increase, on the other hand, for a control group whose mothers were present throughout the observation period (Cox & Campbell, 1968).

3. Autistic children were found to exhibit more stereotyped behavior as the structure of the rooms they were in became more complex. From a series of studies of the frequency of stereotypic behavior in relation to environmental complexity, Hutt and Hutt (1968) developed the hypothesis "that these autistic children were in a state of chronically high cerebral arousal and that stereotypes subserved some mechanism for reducing arousal." "The EEGs of the children were then examined and, surprisingly, were found to consist predominantly of low voltage, and irregular activity with no established rhythms" (Hutt & Hutt, 1970, p. 198). This hypothesis generates the further premise, which is supported also by the investigators' data confirming the well-established clinical observation that autistic children more often fail to make eye-to-eye contact than normals, that failure of eye contact is one means of reducing cortical activity.

Ethologists are studying children's behavior changes under medication and in relation to variations in group density and other social conditions. I shall leave further discussion of what they are doing and finding out to the experts themselves, who are in a much better position to report.

Suffice it to say, the precision brought by the ethologists to the observational measurement of human behavior is most welcome. Some persons would claim that a science does not achieve maturity until

its findings come mainly from laboratory experimentation. I would make the opposite claim for the science of human behavior, namely, that the accuracy of measurement of the phenomena under investigation and the quality of our observations are more fundamental indicators of scientific maturity. Until we have abundant sound information about the full range of behavior under all kinds of natural conditions, we shall not know what is more worthy of experimental manipulation and more controlled research. The relatively atheoretical[1] methods of both the ecologists and ethologists are needed to gather the information necessary for better hypothesis and theory building. The special contribution of ethology to the science of human behavior would seem to be the precision with which observations are made. Such precision is important if naturalistic data are to be considered scientifically trustworthy.

My greatest concern is for the current relevance of ethological work to important educational questions. By focusing so closely on motor patterns and discrete actions, the ethologist might neglect many important areas of human functioning and development in which we are most interested. I hope that ethologists not only will address the methodological questions and issues, but will also indicate the kinds of educational answers we ultimately might expect from them. Ethology may be one of the best hopes for overcoming some of the shortcomings of present day psychology as a scientific discipline, but its direct relevance to educational concerns may be more difficult to see. The practical educator wants to know what he will be able to gain from an ethological approach. If the data are limited to just those items that can be assessed accurately, what do we do about the rest? I suspect that if the research coming from this approach focuses primarily on institutional and preschool children, the transfer of findings and implications to other areas will be minimal. Will the push for precision prevent advances in those variables we consider most important? From each of the other perspectives, there are researchers extending these approaches to the solution of educational problems. Human ethology has yet to prove itself to the practically minded educator. The danger, from the educator's point of view, is that the ethologist will conduct magnificent, tightly designed field studies of children that approach the measurement precision and scientific respectability of experimental research, but the choice of

---

[1]Ethologists claim substantial allegiance to evolutionary theory in what they choose to study and how they interpret their data. Much of their exploratory work, nevertheless, seems inductive and less theory bound than that of most behavioral scientists.

his studies will be irrelevant to the concerns and interests of the educator.

Interaction Analysis

Variations in observational data can be reduced to three general types: narrative, checklist, and ratings. Narrative data, which include specimen records, audio and video tape recordings, and time-lapse photographs and cinematographic records, represent attempts to reproduce faithfully and as completely as possible the totality of events in the fashion and sequence of their original occurrences. The data have the advantage of permitting reflective analysis at a later date using whatever procedures one might wish for examining behavior patterns and relationships. They permit several different dimensions of the same behavior sequences to be studied and possible interrelations among these dimensions to be explored. They make it possible for different investigators to apply their pet theories and methods of analysis to the same material, enhancing the likelihood that overlap among theorists might well be identified and parsimony injected into ultimate explanations. They allow scales and checklists to be perfected to the point where interjudge agreement is high and behavior stabilities readily ascertained.

The problem with narrative records is that they must be broken down in some manner through the use of checklists or rating scales, the other two general types of observational data-gathering tools, before they can become truly useful. Narrative records do not really produce data until they have been coded or rated systematically in some fashion. Thus there are truly only two kinds of observational data to consider: ratings and checklists.

Traditionally, ratings have been the most widely used means of quantifying the observations made by both researchers and practitioners. Unfortunately, the weaknesses of rating scales are many. Naturalistic research has lacked acceptance within behavioral science in no small part because of an overdependence on and misuse of rating procedures for the quantification of observations. Given global and ambiguous trait definitions, halo effects, and other rater sets, to mention only a few of the common deficiencies, ratings too often tell more about the rater than the phenomena which supposedly are under investigation. My biases should not be interpreted as all-condemning, however, as some form of rating is often the best assessment method available for many important human attributes. Quite often, furthermore, higher relations are found between criterion variables and ratings of various treatment or stimulus conditions than categori-

cally defined variables. However, since, along with many others, I have discussed rating procedures more fully elsewhere (Brandt, 1972), let me distinguish quickly between ratings and checklists and move on to the other perspectives.

Ratings represent estimates of the degree to which a particular characteristic is manifest along a presumed continuum, generally ranging from complete absence to full presence. Checklists are sets of predefined categories for classifying and tallying live behavior or narrative records along with relevant contextual features. Observer evaluation is minimal, representing primarily a qualitative judgment that an action or feature is of one category type or another. A teacher verbalization, for example, might be classified as a statement, question, or command, among other possibilities.

All observational methodologists have to be concerned with categorical judgments, but the interactionists in our group probably have the most to say about categories. As a group they have been particularly concerned with the development of category sets to describe crucial variations in human functioning. Their scales make possible the coding and counting of behaviors, events, and interaction sequences with respect to a target person and the persons and objects with which he interacts. The interactionist is interested in pure description of the nature of ongoing events, as are both the ecologist and ethologist. He is also interested in discovering relationships that often exist among the several forces making up ongoing events and in the interactional patterns that are most highly correlated with certain outcomes.

In contrast to the previously discussed perspectives, interactionist observations are theoretically based. Whatever category sets he decides to use are selected from many possibilities. He selects categories that not only are distinguishable one from another but also reflect the dimensions of human activity he considers most relevant to the problem he is exploring. His answers to whatever questions he explores will be limited, of course, to the particular category sets that he uses to record data. His answers will be limited to the frequencies or durations of whatever types of behavior he classifies, to their sequential patterns, and to the interrelation of behavior types.

The interactionist understands the complexity of behavior in naturalistic settings and realizes that he can record fully and accurately only a fraction of all that action. The utility of what he does choose to record depends not only on his systematic sampling of time and events, but also on how thoroughly the categories cover the variables he has presumed to study. Theory is essential, therefore, in isolating

important variables to be investigated, deriving meaningful categories for those variables, and providing a rationale for later interpretation.

Role theory is useful, for example, in developing or selecting a category set for analyzing the nature of group structure and the contributions of individual members to the group. It may be useful also to break down and assess the specific nature of one's job. Reinforcement theory is most helpful, as our behavior modification experts will certainly show, in investigations of both skill learning and planned change in naturalistic settings. Psychoanalytic theory and rationale from depth psychology can provide category sets that cover defense mechanisms and other behavioral manifestations of inner processes. Self-theory can stimulate the types of behavior to be coded in observing loosely structured situations in which individuals have opportunities to choose freely among alternatives and generate their own directions. What promise such category sets hold for improved methods of attitude and interest measurement, even for anxiety and drive!

Checklists come in many varieties. Elsewhere (Brandt, 1972), I have listed at least a dozen somewhat discrete types, such as *activity logs*, *[illegible] records*, and *standardized situation responses*. Most recording requires *static descriptors* to be noted, indicating certain relatively stable, descriptive characteristics of the research subjects (age, sex, code number, etc.) and setting (time, place, activity period, etc.). The largest generic type I refer to is *action checklists*.

Medley and Mitzel (1963, pp. 298-303) would divide action checklists, furthermore, into two kinds: (a) *category systems* and (b) *sign systems*. The well-known Flanders interaction system illustrates the former, with each observed behavioral unit classified into one and only one category. Each category in the set must be mutually exclusive and independent of each other category and the set as a whole must be exhaustive, that is, a specific behavioral unit must always be classifiable somewhere in the set, which might need to include a "miscellaneous" category. With a sign system, a number of discrete behaviors are precisely identified and noted should they occur. At the end of a specified time interval (10 seconds, 2 minutes, etc.), whichever of the behaviors has occurred (from none to all) is noted on the recording form.

Compared with a category system, a sign system is usually made up of a much larger number of types of behavior to be observed, but the behavior types are more narrowly defined and occur much less frequently. Whereas an observer using a category system must record every behavioral or time unit, one employing a sign system may watch for relatively long periods without seeing and recording any behaviors,

simply because the particular behaviors making up the sign system do not occur (Brandt, 1972, p. 102).

Examples of teacher behavior that might go into a sign system because they occur so infrequently that they are not likely to show up at all on a time-sampling category schedule are "threatens a child," "sends a child to the office," "cries," "swears," "calls a parent," and the like. Using a sign system that would include such a list of specific but infrequently occurring behaviors, an observer might even watch for and check any of these behaviors if they occur while he is using a regular category system.

One other basic distinction is when the recording is made. For some observation systems behavior is coded on some pre-specified, *systematic time basis* (every 3 seconds, etc.). For *event sampling*, on the other hand, behavior events are recorded in a specified manner whenever they occur. Both the time the act happened and its duration may be part of the recorded data even though the basic schedule is not based on the sampling of time. It is often possible, therefore, to interpret event-sampling as well as time-sampling data on a frequency per unit of time basis. In the Follower-Leader scale that my students are constructing, for example, the numbers of specific acts of either type per unit of time can be reported if the observer merely records the initiation and termination times for his observation periods. One advantage of event sampling is that the several components making up the event may be coded as well as the mere existence of the event. With our Leader-Follower scale, for example, the sequence of the interaction, the persons involved, and the specific type of leader-follower behavior are coded sequentially. I am sure that the various presenters will be citing numerous illustrations of both time and event sampling.

Action checklists differ also with respect to the molarity of the action units being covered. A log might be designed for recording selected characteristics of the major shifts in instructional activity over the school day. Entries would be made only when the overall organization and instruction patterns shifted from one type of activity to another, perhaps only two or three times an hour. More generally, however, interaction schedules are focused on individual but still molar behaviors, such as "asks a question" or "gives a direction," or on such molecular behavior as would interest the ethologist, like "smiles" or "frowns."

In addition to the variations noted in the preceding discussion, one can distinguish between general observation systems, such as Medley's OScAR and PROSE instruments, and specific observational scales, like Flanders interaction scale, in which only one dimension is being

observed. Made up of a number of separate categorical sets and perhaps including sign sets as well, general observation systems permit an observer to code a number of important dimensions of classroom life at the same time, thus approaching the totality of coverage that the ecologist sees but in a more precise manner.

The most extensive usage of such a system is probably the Stanford Research Institute's ongoing evaluation study of Follow Through programs. Stallings and her colleagues (Stanford Research Institute, 1972) have developed a multidimensional schedule covering dozens of variables which were selected by Follow Through sponsors as important descriptors of their programs. One section requires an observer to code, at the end of a school visit, certain relatively stable features of the physical environment, such as the amount of playground space and air-conditioning and heating provisions. In another section, the observer records in snapshot fashion every 15 minutes where everybody in the room is, what each adult and child is doing, the size of groups, and the types of activities in progress. In the third section, the teacher or a child is followed for five-minute periods four times an hour; interaction sequences are recorded on four types of categories providing data on "Who does the action?" "To whom is it done?" "What is done?" and "How is it done?"

This instrument is proving to be quite useful in determining how well a program is implemented at particular sites. As indicated earlier, it is also providing considerable information on instructional process-child outcome relationships. In addition, it allows program comparisons to be made both within and between various models in order to discover the specific points of difference and similarity, despite whatever theoretical claims might be made.

Stallings (1973) has reported, for example, that the programs of the Universities of Oregon and Kansas typically yield higher interactional frequencies on such items as (a) an adult informing a child, (b) an adult asking a child a direct question, (c) an adult direct question being followed by a child response, (d) a child response being followed by adult feedback, and (e) academic activities in progress. The Educational Development Center's program, among other open education models, is characterized by relatively high frequencies of teachers asking open-ended questions.

One common weakness of comprehensive observation systems is their tendency to be somewhat content free; they focus directly on process description. A second weakness is the rather limited specificity of the categories that relate to a given dimension. If a single observer must code a number of dimensions, the number of categories per dimension is often as few as two and seldom greater than six or seven.

Single dimension scales need to be used, therefore, to generate more precise data about more narrowly focused types of behavior (e.g., reading style or different types of social cooperation). In designing and using single dimension scales, one must consider the kinds of situations in which they are most applicable, that is, situations in which behavioral data relevant to the particular dimensions being assessed can be obtained most readily. Furthermore, using such scales in a variety of specified situations helps not only to discern the impact of contextual variables but also to assess the stability of particular behavior patterns. Natchez (1959), for example, found that the type of defensive behavior displayed by poor readers during daily oral reading activities was highly consistent with that evidenced on the playground and also noted by teachers in their report card comments.

Observation checklists, therefore, take a variety of forms and serve a number of purposes. They vary sharply in coverage and the target object for investigation. They can be designed for adult or child observations; for content exposure or task analysis; for activities and events, for interaction dyads, and the like. Although arguments can be made for conducting research on each of these observation targets, educational research in the past has been concentrated primarily on teachers.

Top priority should be given to children's school behavior and to their content exposure and task demands, especially. Little is known about within-class variability of children's experience and performance in relation to specific content and task requirements. Most child research efforts have attempted to determine grade-level, sex, ethnic, socioeconomic, or ability differences rather than the idiographic patterns of child response or, more precisely, their behavioral responses to differing instructional conditions and demands. I would hope that we can give some attention at this conference to the problems inherent in within-child observational studies and research which focuses primarily on the child rather than her/his teacher, the classroom as a whole, or a model method or program as the target of analysis.

One reason the structured Follow Through models may be showing up as more successful in teaching academic skills is the likelihood that they give children significantly greater practice and content exposure than more open programs do. In the latter, the program is highly individualized and, unless records are kept on a child-by-child basis, we cannot know how important this factor may be.

Almost no research has been devoted to assessing specific task demands and children's responses to various instructional materials

despite the great proliferation of curricula programs. Some assignments call for divergent responses and others convergent, to mention only one dimension of difference among them. The dyadic relationships with people—adults and other children—likewise are seldom recorded. Thus, sequential stimulus-response types of analysis cannot be made. Rather, most child observation studies to date have merely recorded frequencies of child behavior by categories.

Behavior Modification

Let me proceed to a discussion of the final perspective, behavior modification. More than any of the other perspectives, observation studies by behavior modification experts have focused on individual children. However, the scope of their study is typically very narrow; only a piece of the whole child is targeted for observation (e.g., his stuttering, inattention to task, talking out, etc.). Concentration on individual target behaviors for modification, furthermore, has even brought forth a new type of research design. Experimental control no longer rests in matched comparison or experimental vs. control groups but in within-subject changes from baseline rates of target behavior during treatment and reversal-of-treatment conditions.

A single theoretical perspective underlies the entire behavior modification movement based essentially on operant conditioning. The historical roots of the movement appear in the works of Thorndike, Hull, Miller and Dollard, and especially Skinner. The fundamental learning principle underlying behavior modification techniques is that reinforced behavior tends to persist. Behavioral change, therefore, is accomplished by ignoring undesirable behaviors when they occur and reinforcing desirable ones.

Observation is the primary vehicle for determining what behavior patterns exist at various stages in a program and what type or amount of reinforcement the natural environment provides. Observational data are also collected to monitor the implementation of a behavior modification program and to identify personal preferences among natural reinforcers for use in the programs. Under the "Premack principle," high-probability behavior can be used as a natural reinforcer for low-probability behavior. If a child watches television a good deal and seldom studies his school work, parents might permit television watching only as a reward for an increased amount of studying.

Much can be learned about ongoing school life from close inspection of reinforcement patterns as well as the kinds of behavior that are and are not being reinforced. Despite the clarity of educational

planning and the best of teacher intentions, schools teach well many undesirable learnings and poorly many desirable learnings, through unintentional reinforcement. Because of the complexity of school life, it is only through systematic observation that such inappropriate teaching patterns can be recognized and corrected.

Many illuminating studies, not overly demanding in terms of resources, are possible on the part of one observer—a school supervisor, psychologist, or the teacher himself—for assessing current reinforcement patterns. I have cited examples elsewhere (Brandt, 1972) of studies of (a) teacher evaluation criteria that are reflected in the remarks and symbols placed on student compositions as they are graded; (b) instructional emphases in terms of intellectual and non-intellectual attributes and products, operations, and content, using Guilford's model, that are evident in records kept of class assignments, and any supplementary oral directions and specifications teachers might give for completing the assignment; (c) differential teacher enforcement of classroom rules among pupils; and (d) teacher reinforcement patterns in relation to specific types of expectancies and to individual children. These studies do not exemplify the behavior modification model as directly as some of the child-change studies that have been reported by others at this conference, but they do show how observation can be used to assess the types of reinforcements the school is providing. They could easily serve as models for examining current practices, as a basis for planned change, and for indicating the kinds of measures that might be used in assessing the success of change attempts.

Behavior modification records show the frequencies or time rates of particular kinds of behavior. These records can show quite precisely how a child is performing in selected instructional areas. They often serve diagnostically to help teachers match the materials and expectancies of instruction to a child's performance or ability level. In general, these records are probably the best available of the child's specific learning accomplishments and deficiencies. They offer a means for continually diagnosing instructional needs and sequencing instructional demands.

One should be aware, however, that behavior modification records do not usually show how learning might transfer to other situations, what changes in cognitive functioning and structure might be occurring, or what possible detrimental side effects also might be taking place. The specificity of the records highly restricts the focus to those target behaviors under investigation.

Nevertheless, numerous successes have been reported in regard to both specific learning skills and, in some instances, more general

measures of achievement. The structured Follow Through programs have tended to show greater gains in arithmetic and language areas than most other programs, perhaps, in considerable part, because of the extra time spent on skill development in these areas. The specific adult question, child answer, and adult feedback pattern used so frequently in these programs demonstrate considerable power as an instructional procedure. That the children learn to be more task persistent is apparent too, but they may also learn to be less self-sufficient and creative. This observation is not a firm criticism on my part, only the basis of an open question that might be raised.

I have strayed considerably from the straight behavior modification model to consider curricular programs based on behavioristic psychology and studies of school reinforcement practices specifically; but I do so because I see them as interrelated. The behaviorists' successes are often disarming to the non-behaviorist. They have not solved nor are they likely to solve all of our educational problems. Their push for specificity and precision is also their undoing, in the larger sense, because there is little hope that the hundreds of specific response skills and knowledge which they might be able to build into an educational program will ever approach the much larger array of learning and knowledge needed by the developing child. Transfer is the critical question and, if the behavior modification techniques do not produce transfer on a fairly broad scale nor promote wholesome incidental learnings—as yet to be determined—the approach will remain limited in its utility and only one tool among many.

Some Promises and Issues

During the first half of the twentieth century, the research support for learning theory and curriculum development was fragmentary and oversimplistic. Despite the monumental contributions of Thorndike and the magnificent attempt at a comprehensive evaluation of modern curriculum development represented by the Eight Year Study, education lacked the empirical base that has characterized medicine, engineering, and even the legal profession.

In attempting to establish such a base, human development and learning theorists overgeneralized the findings of small-scale research studies, which were often conducted in laboratory settings, to draw major implications for school practice. One example of such overapplication of basic research to educational policy was justifying the superiority of pupil-teacher planning, discussion-oriented instruction, and democratic classroom climate primarily from the results of the

Lewin, Lippitt, and White (1939) studies of the behavior and productivity of boys in small social clubs. The prevailing implications drawn from this research were that teachers should not be authoritarian, establish firm expectancies, structure assignments tightly, nor evaluate youngsters too rigorously if they want children to remain productive, intrinsically motivated, and psychologically undamaged. It was not until Flanders (1961) and his associates employed naturalistic observation methods to assess teaching style directly that substantial empirical evidence was obtained to indicate that it was not democracy but teacher flexibility that led to superior learning and improved attitude. The "good teacher" was at times highly directive, but at other times (s)he shifted roles and permitted children considerable self-direction and high activity responses.

There is now a growing recognition of the importance and practicality of empirical research which is conducted directly on ongoing school operations. Solid data systematically gathered on various aspects of one's own program can provide a sound basis for educational decisions without taking the gigantic and often erroneous cognitive leap from irrelevant, single variable, laboratory studies to recommendations for school practice.

Despite many common threads, the uniqueness of each school, the population it serves, and the personnel resources it commands suggest that such studies need to be done routinely within the particular school's own walls rather then depend on large-scale research answers that are supposedly applicable to all institutions. Because of the interactive complexity of the many variables affecting particular educational outcomes, it has always been too easy to dismiss research findings reported elsewhere as not applicable to a given institution. When the data are derived locally, however, greater credibility of findings cannot help but have more profound impact on practice.

Successful accomplishment of such studies will depend primarily on local school staffs (school psychologists, counselors, curriculum specialists, supervisors, project directors, and teachers in some instances) understanding the developing technology and having sufficient administrative encouragement to employ it. In many instances, the research literature today offers school research practitioners the specific observational methods and scales that are needed for many of the studies they might wish to conduct. Where technology is lacking, furthermore, considerable help for developing particular instruments and research plans can be obtained from several recent books (Beegle & Brandt, 1973; Boehm & Weinberg, in press; Brandt, 1972; Good & Brophy, 1973; Griffith, 1973; Hall, 1973; Hunkins, 1972; Kounin, 1970; and Tharp & Wetzel, 1969).

Although individual priorities will need to be set for deciding the order and manner in which studies are to be done, and the conduct of studies will depend on clarifying purposes rather specifically before tools and procedures can be selected with finality, several kinds of studies would seem particularly needed.

1. *Behavioral surveys* of what teachers, administrators, and pupils do in the course of the school day represent one kind of study needed—how much various individuals talk, listen, read, watch, and so forth; what kinds of roles they assume; and what kind of experiences they truly have. Fleming (1973) reported one such shadow study in which he found that secondary teachers engaged in almost constant talking, regardless of their subject, and that students had almost no opportunity to raise questions, contribute to discussion, or do anything except sit and listen. Jackson (1968) kept records that showed that shocking amounts of student time were taken up in waiting to make active responses and that student activities were frequently interrupted by bells, teacher directions, schoolwide announcements on the PA system—factors that would hardly promote learning.

2. *Program monitoring*, described earlier, is a second general type of in-school study. Records are useful to show the extent to which key aspects of planned programs are, in fact, being implemented. Monitoring should include not only instructional processes but content coverage and analysis as well.

3. Another type of in-school research needed might be labeled *situational response studies*. Many opportunities exist in the ordinary structuring of classroom activity and administration of the school program to measure student learning merely by recording their responses to either contrived or naturally occurring situations without increasing the amount of testing done. Silent reading rates can be assessed by recording how many pages are covered in particular books during a given period of uninterrupted time. The extent of black-white integration can be partially determined, at least as one of our graduate students did, by merely recording who was sitting beside whom on the bus and in the cafeteria. The effectiveness of cafeteria clean-up campaigns can be assessed by the amount of trash left after lunch is over. The relative effectiveness of two signs in keeping students on the sidewalk and off the shortcut across the lawn was very apparent on our campus recently. When a traditional "Stay off the Grass" sign was replaced by one reading "Ecology Please," the worn path soon disappeared. Although one must be systematic and scientifically sound in his use of unobtrusive monitoring, as well as cautious in how much he can generalize from one

type of assessment by itself, the opportunity to make such studies is only limited by our imaginations.

4. *Case studies of individual pupils* especially are needed now, as they always have been needed, to assess developmental change, understand more clearly the forces that shape behavior, and recognize fully the differential nature of school life for the many youngsters served by the school. The scales and observational techniques now available make it possible to conduct much better case studies than in the past. Almost no one seems to be studying individual children with a comprehensive assortment of behavioral scales, but the possibilities exist for solid, trustworthy school records to be kept of individual children from which we should learn much.

By itself, study of an individual is too narrow a base for total system change. Perhaps the most important function served by detailed analysis of student and teacher behavior is the demonstration to educators of the complex nature of all aspects of school life. Full acceptance of that complexity cannot help but stimulate reconsideration of program objectives and improve curriculum development.

Without objective data, I am inclined to think that we fool ourselves often in thinking that change and improved practice have occurred because teachers or administrators talk the latest educational jargon well. As I have indicated elsewhere (Brandt, 1972), Goodlad, Klein, and their associates (1970), in their excellent observational study of 158 classes in 67 schools, showed that,

> (1) Many widely recommended educational improvements were not really taking place. (2) Even though many teachers thought they were providing individualized instruction, encouraging inductive learning, and using group dynamics principles, observable evidence that these innovations were being practiced was seldom found. (3) "Special" supplementary and enrichment activities differed very little from ordinary class activities. (4) Classroom goals were usually not identifiable to observers; and instruction was seldom directed toward the diagnosed needs, progress, and problems of individual children. In brief, classroom instruction was being conducted along very traditional patterns.
>
> Until a considerable amount of purely descriptive investigation occurs, one can only guess at how widespread particular practices are or how likely the findings of selected studies are to apply elsewhere. (Brandt, 1972, p. 355)

One of the purposes in bringing together experts representing rather diverse observational methodologies is to raise issues and attempt to resolve some of the seeming discrepancies among them

with respect to what is important to study and how one might best go about in-school research. Sharp differences would seem to exist with respect to the target of investigation, investigational focus, variables selected for study, the type of data collected, the size of behavioral units, sampling procedures, and the overall scope of investigations. Differences exist also in the way data are processed and treated statistically, as a basis for interpreting findings. How best to determine the reliability and validity of observational measures is certainly worthy of debate and resolution.

While vast differences most certainly do exist in the general purposes and the specific methodologies of our experts, some points of agreement would seem apparent at the outset. One is the concern for objective data that accurately describe whatever is being reported. There is agreement also in the importance of studying ongoing activity under real-life conditions.

A major problem presented by objective reporting, whatever its form, is what Thorndike and Hagen (1961) label "outsideness." Behavior is but the outward manifestation of attitudes and other subjective components of human functioning. Objective recording cannot directly assess these components. Internal cognitive and affective dimensions, therefore, present special measurement problems for the behavioral scientist generally, so the combined use of both direct observation and other types of data-gathering tools (e.g., questionnaire tests, interviews) must also be considered.

The power of observational tools and techniques has barely begun to be realized. By sharing our thinking at this conference, I hope that we will not only clarify the similarities and the differences among the approaches and sharpen educators' awareness of the applicability of each of them, but that we will resolve as well some of the issues that prevent more widespread usage of observational techniques.

References

Barker, R.G. *Ecological psychology: Concepts and methods for studying the environment of human behavior.* Stanford, Calif.: Stanford University, 1968.

Barker, R.G., & Gump, P.V. *Big school, small school: High school size and student behavior.* Stanford, Calif.: Stanford University, 1964.

Barker, R.G., & Wright, H.F. *Midwest and its children: The psychological ecology of an American town.* N.Y.: Harper & Row, 1955.

Beegle, C.W., & Brandt, R.M. (Eds.) *Observational methods in the classroom.* Washington, D.C.: Association for Supervision and Curriculum Development, 1973.

Boehm, A.E., & Weinberg, R.A. *The classroom observer: A guide for developing observational skills*. N.Y.: Teachers College, in press.

Brandt, R.M. *Studying behavior in natural settings*. N.Y.: Holt, Rinehart & Winston, 1972.

Brandt, R.M. Three weeks in British infant schools. In D.D. Hearn, J. Burdin, & L. Katz (Eds.), *Current research and perspectives in open education*. Washington, D.C.: American Association of Elementary-Kindergarten-Nursery Educators, 1973.

Carthy, J.D. *The study of behavior*. London: Arnold, 1966.

Cox, F.N., & Campbell, D. Young children in a new situation with and without their mothers. *Child Development*, 1968, **39**, 123-131.

Flanders, N.A. Analyzing teacher behavior. *Educational Leadership*, 1961, **19**, 173-180.

Flanders, N.A. *Analyzing teacher behavior*. Reading, Mass.: Addison-Wesley, 1970.

Fleming, R.S. The supervisor as an observer. In C.W. Beegle & R.M. Brandt (Eds.), *Observational methods in the classroom*. Washington, D.C.: Association for Supervision and Curriculum Development, 1973.

Goldupp, O Observational findings regarding independent learning behavior. Paper presented at the meeting of the American Educational Research Association, Chicago, April 1974.

Good, T.L., & Brophy, J.E. *Looking in classrooms*. N.Y.: Harper & Row, 1973.

Goodlad, J.I., Klein, M.F., & associates. *Behind the classroom door*. Worthington, Ohio: Charles A. Jones, 1970.

Griffith, F. *A handbook for the observation of teaching and learning*. Midland, Mich.: Pendell Publishing, 1973.

Gutmann, D. Psychological naturalism in cross cultural studies. In E.P. Willems & H.L. Raush (Eds.), *Naturalistic viewpoints in psychological research*. N.Y.: Holt, Rinehart & Winston, 1969.

Hall, R.V. *Managing behavior*. Part 1, Behavior modification: The measurement of behavior. Part 2, Behavior modification: Basic principles. Part 3, Behavior modification: Applications in school and home. Lawrence, Kan.: H & H Enterprises, 1973.

Hunkins, F.P. *Questioning strategies and techniques*. Boston: Allyn & Bacon, 1972.

Hutt, S.J., & Hutt, C.H. Stereotypy arousal and autism. *Human Development*, 1968, *11*, 277-286.

Hutt, S.J., & Hutt, C.H. *Direct observation and measurement of behavior*. Springfield, Ill.: Charles C. Thomas, 1970.

Jackson, P.W. *Life in classrooms*. N.Y.: Holt, Rinehart & Winston, 1968.

Kounin, J.S. *Discipline and group management in classrooms*. N.Y. Holt, Rinehart & Winston, 1970.

Lewin, K., Lippitt, R., White, R.K. Patterns of aggressive behavior in experimentally created "social climates." *Journal of Social Psychology*, 1939, **10**, 271-300.

McGrew, W.C. Glossary of motor patterns of four-year-old nursery school children. In S.J. Hutt & C.H. Hutt, *Direct observation and measurement of behavior*. Springfield, Ill.: Charles C. Thomas, 1970, pp. 210-218 (Appendix).

McGrew, W.C. *An ethological study of children's behavior*. N.Y.: Academic Press, 1972.

Medley, D.M. Closing the gap between research in teacher effectiveness and the teacher education curriculum. *Journal of Research and Development in Education*, 1973, **7**, 39-46.

Medley, D.M., & Mitzel, H.E. Measuring classroom behavior by systematic observation. In N. Gage (Ed.), *Handbook of research on teaching*. Chicago: Rand McNally, 1963.

Natchez, G. *Personality patterns and oral reading*. N.Y.: New York University, 1959.

Ross, S., & Zimiles, H. Children's interactions in Follow Through classrooms: The DCB observation system. Paper presented at the meeting of the American Educational Research Association, Chicago, April 1974.

Schoggen, M., Barker, L.S., & Barker, R.G. Structure of the behavior of American and English children. In R.G. Barker (Ed.), *The stream of behavior*. N.Y.: Appleton, 1963.

Sechrest, L. Nonreactive assessment of attitudes. In E.P. Willems & H.L. Raush (Eds.), *Naturalistic viewpoints in psychological research*. N.Y.: Holt, Rinehart & Winston, 1969.

Simon, A., & Boyer, E.G. (Eds.) *Mirrors for behavior: An anthology of classroom observation instruments*. Vols. 1-6. Philadelphia: Research for Better Schools, 1967.

Simon, A., & Boyer, E.G. (Eds). *Mirrors for behavior: An anthology of classroom observation instruments*. Vols. 7-14 and Summary. Philadelphia: Research for Better Schools, 1970. (a)

Simon, A., & Boyer, E.G. (Eds.) *Mirrors for behavior: An anthology of classroom observation instruments*. Supplementary Vols. A and B. Philadelphia: Research for Better Schools, 1970. (b)

Smith, L.M., & Geoffrey, W. *The complexities of an urban classroom*. N.Y.: Holt, Rinehart & Winston, 1968.

Soar, R.S. An empirical analysis of selected Follow Through programs: An example of a process approach to evaluation. In 1971 *Yearbook for early childhood education*. Chicago: National Society for the Study of Education, 1972. (a)

Soar, R.S. Teacher-pupil interaction. In *A new look at progressive education*, 1972 Yearbook. Washington, D.C.: Association for Supervision and Curriculum Development, 1972. (b)

Stallings, J. Follow Through program classroom observation evaluation, 1971-72. (Final report for Bureau of Elementary and Secondary Education, U.S. Office of Education.) Menlo Park, Calif.: Stanford Research Institute, 1973.

Stallings, J. Is compensatory education working? Paper presented at the meeting of the American Association of Elementary-Kindergarten-Nursery Educators Research Committee Conference, January 1974. (a)

Stallings, J. An implementation study of seven Follow Through models for education. Paper presented at the annual conference of the American Educational Research Association, Chicago, April 1974. (b)

Stanford Research Institute. *Training manual for classroom observation*. Menlo Park, Calif.: Stanford Research Institute, January 1972.

Stearns, M.S. *Report on preschool programs: The effects on disadvantaged children and their families*. Washington, D.C.: Office of Child Development, U.S. Department of Health, Education, and Welfare, 1971.

Tharp, R.G. & Wetzel, R.J. *Behavior modification in the natural environment*. N.Y.: Academic Press, 1969.

Thorndike, R.L. & Hagen, E. *Measurement and evaluation in psychology and education*. (2nd ed.) N.Y.: Wiley, 1961.

Willems, E.P., & Raush, H.L. (Eds.) *Naturalistic viewpoints in psychological research*. N.Y.: Holt, Rinehart & Winston, 1969.

Wright, H.F. Observational child study. In P.H. Mussen (Ed.), *Handbook of research methods in child development*. N.Y.: Wiley, 1960.

Wright, H.R. *Recording and analyzing child behavior*. N.Y.: Harper & Row, 1967.

"END OF DOCUMENT"